

PA (HARD) HARVARD COLLEGE.
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
XX
XX Ingham PW, McMahon AP, Tablin CJ;
XX WPI, 1995-255060/33.
DR N-PSDB; AAO91636.
XX
XX Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
PT to treat degenerative nervous system disorder(s) and in gene
PT therapy.
XX
XX Claim 17; Page 133-35; 210pp; English.
XX
XX The sequence represents a chicken sonic hedgehog protein,
CC homologous to a Drosophila hedgehog protein (AAR77337), and is
CC encoded by a cDNA isolated from a stage 22/22 limb bud cDNA
CC library. Probes and primers derived from the sonic hedgehog gene
CC may be used as diagnostic agents for neuromuscular, autonomic or
CC central nervous system disorders, and the gene may also be used in
CC gene therapy. Antibodies generated from the protein may be used
CC as therapeutic or research reagents.
XX
XX
SQ Sequence 425 AA;
Query Match 100.0%; Score 2218; DB 16; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.7e-227;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WEMLLRILLYGFTICALLVSSGLTCGPGRGIGRRHPRKKTPLAYKQFINVAEKLIG 60
DB 1 WEMLLRILLYGFTICALLVSSGLTCGPGRGIGRRHPRKKTPLAYKQFINVAEKLIG 60
OY 61 ASGREGKITRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALATISVMNOM 120
DB 61 ASGREGKITRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALATISVMNOM 120
OY 121 PCVKLRVTEGMDGDHSEESLHYEGRAVDITTSRDRSKYGMRLAVEAGFDWVYYES 180
DB 121 PCVKLRVTEGMDGDHSEESLHYEGRAVDITTSRDRSKYGMRLAVEAGFDWVYYES 180
OY 181 KAHHCYVAKENSVAKSGCGPGSATVHLEHGTRKIVKDSPGDRVLAADADGRILYSD 240
DB 181 KAHHCYVAKENSVAKSGCGPGSATVHLEHGTRKIVKDSPGDRVLAADADGRILYSD 240
OY 241 FLTFLDRMDSRKLFYIETROPARALLTAHLLFVAPOHNSBATGSGTSGOALFASNV 300
DB 241 FLTFLDRMDSRKLFYIETROPARALLTAHLLFVAPOHNSBATGSGTSGOALFASNV 300
OY 301 KPGQVYVVLGEGGQULLPASVHVSGLREBASGAYAPLTAOGTILINRYLASCYAVIEEHS 360
DB 301 KPGQVYVVLGEGGQULLPASVHVSGLREBASGAYAPLTAOGTILINRYLASCYAVIEEHS 360
OY 361 WAHMAFAPRLAOGLLAALCPDGAIPTAATTTGTHWYSRLLYRIGSWLQSDALHPLGM 420
DB 361 WAHMAFAPRLAOGLLAALCPDGAIPTAATTTGTHWYSRLLYRIGSWLQSDALHPLGM 420
OY 421 VAPAS 425
DB 421 VAPAS 425
RESULT 2
ID AAY05854 standard; Protein: 425 AA.
XX
XX AAY05854;
XX
XX 02-AUG-1999 (first entry)
XX
XX Chicken Sonic hedgehog Shh protein.
XX
XX Sonic hedgehog; Shh; chicken; epithelial tissue; epithelium;

KM cutaneous tissue; skin; hair; wound healing; vulnerrary;
KM burn; skin grafting; pressure sore; ulcer; ulcerative colitis;
KM alopecia; psoriasis; keratosis; acne; comedogenic lesion;
KM folliculitis; pseudofolliculitis; keratoacanthoma; callositles;
KM Darier's disease; scar; autoimmune disease; pemphigus;
KM epidermolysis; lupus lesion; desquamative lesion; carcinoma;
KM therapy; hedgehog therapeutic; plc therapeutic; patched.
XX
XX Gallus domesticus.
XX
XX WO9920298-A1.
XX
XX 29-APR-1999.
XX
XX 20-OCT-1998; 98WO-US22227.
XX
XX 11-SEP-1998; 98US-0151999.
XX 20-OCT-1997; 97US-0955552.
XX
XX (OMTO-) ONTOGENY INC.
XX
XX Wang EA;
XX
XX WPI, 1999-288170/24.
XX N-PSDB; AAX25617.
XX
XX Use of hedgehog polypeptides on patched therapeutics
XX
XX Claim 26; Page 119-120; 146pp; English.
XX
XX The present sequence represents chicken Sonic hedgehog protein
CC Shh. The invention relates to a method for modulating the growth
CC state an epithelial cell by ecotopically contacting the epithelial
CC cell, in vitro or in vivo, with a hedgehog therapeutic (i.e. a
CC hedgehog polypeptide or gene therapy construct) or ptc therapeutic
CC (i.e. a small organic molecule that mimics the effect of hedgehog
CC proteins on patched signalling, or activates or potentiates patched
CC signalling) in an amount effective to alter the rate of proliferation
CC of the epithelial cell. The hedgehog therapeutic preferably
CC comprises at least a bioactive extracellular portion of a hedgehog
CC protein (see AAY05854-62) encoded by a vertebrate hedgehog gene (see
CC AAX25617-25), especially a human hedgehog gene. Promotion of
CC proliferation of epithelial cells can be used to control a wound
CC healing process in e.g. burn treatment, skin regeneration, skin
CC grafting, pressure sore treatment, dermal ulcer treatment, post
CC surgery scar reduction or treatment of ulcerative colitis (claimed).
CC It can also be used to induce hair growth for the treatment of
CC alopecia (claimed). Inhibition of the growth of epithelial tissue
CC can be used to treat or prevent hyperplastic or neoplastic
CC conditions, e.g. psoriasis, keratosis, acne, comedogenic lesions,
CC folliculitis and pseudofolliculitis, keratoacanthoma, callositles,
CC Darier's disease, keloids, hypertrophic scars, or autoimmune
CC disorders, e.g. aphthous ulcers, pemphigus vulgaris, pemphigus
CC foliaceus, pemphigus vegetans, pemphigus erythematous,
CC epidermolysis, lupus lesions, desquamative lesions or carcinomas.
CC The methods can also be used to counteract the effects of aging on
CC skin.
XX
XX
SQ Sequence 425 AA;
Query Match 100.0%; Score 2218; DB 20; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.7e-227;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WEMLLRILLYGFTICALLVSSGLTCGPGRGIGRRHPRKKTPLAYKQFINVAEKLIG 60
DB 1 WEMLLRILLYGFTICALLVSSGLTCGPGRGIGRRHPRKKTPLAYKQFINVAEKLIG 60
OY 61 ASGREGKITRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALATISVMNOM 120
DB 61 ASGREGKITRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALATISVMNOM 120
OY 121 PCVKLRVTEGMDGDHSEESLHYEGRAVDITTSRDRSKYGMRLAVEAGFDWVYYES 180

|||||
Db 121 PGVKLRVTEGMDGDHSESLHTEGRAVDITTSDDRSKYGLARLAVAGFDWYYES 180
Oy 181 KAHHCSVKAENSVAAKSGCFPGSATVHLEHGCTKLVKDLSFGDRVLAADADGRLLYSD 240
Db 181 KAHHCSVKAENSVAAKSGCFPGSATVHLEHGCTKLVKDLSFGDRVLAADADGRLLYSD 240
Oy 241 FLTFIDRMDSSRKLFFYIETROPARALLTAHLLFVAPDHNSGATGSGALFASNV 300
Db 241 FLTFIDRMDSSRKLFFYIETROPARALLTAHLLFVAPDHNSGATGSGALFASNV 300
Oy 301 KEGQRYVYGEQGQQLLPASVHSVSLREASGAYAPLTAOGTLLINRVLASCAVYEHS 360
Db 301 KEGQRYVYGEQGQQLLPASVHSVSLREASGAYAPLTAOGTLLINRVLASCAVYEHS 360
Oy 361 WAHMAFAPFRLAOGLLAALCPDGAIPTAATTTGIMHYSRLYRIGSWLDGALPLGM 420
Db 361 WAHMAFAPFRLAOGLLAALCPDGAIPTAATTTGIMHYSRLYRIGSWLDGALPLGM 420
Oy 421 VAPAS 425
Db 421 VAPAS 425
RESULT 3
AAV05510
ID AAV05510 standard; Protein; 425 AA.
AC AAV05510;
XX
XX 05-JUL-1999 (first entry)
DE Chicken Sonic hedgehog protein Shh.
XX
XX Sonic hedgehog; Shh protein; chicken; hedgehog therapeutic;
KW ptc therapeutic; patched; signal transduction; muscle atrophy;
KW cachexia; muscular myopathy; myoblastic sarcoma; therapy.
OS Gallus sp.
XX
XX MO9910004-A2.
PN
XX 04-MAR-1999.
PD
XX 28-AUG-1998; 98WO-US17922.
PF
XX 29-AUG-1997; 97US-0057394.
PR
XX (ONTO-) ONTOGENY INC.
PA
XX Bladgen CS, Currie PD, Hughes SM, Ingham PW;
PI MPI: 1999-243557/20.
XX DR N-PSDB; AAX25098.
XX
XX A new method to regulate muscle growth
PT
XX
XX Disclosure: Page 111-112; 130pp; English.
PS
XX
XX The present sequence is chicken Sonic hedgehog protein Shh. The
CC invention relates to a method for modulating the formation and/or
CC maintenance of muscle tissue by ecotopically contacting muscle
CC cells, especially muscle stem/progenitor cells, in vitro or in
CC vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and
CC gene therapy constructs) or ptc therapeutic (i.e. a small organic
CC molecule that mimics the effect of hedgehog proteins on patched
CC signaling, or activators or potentiators patched signaling) in an
CC amount effective to alter the growth state of the treated cells.
CC Also claimed is a method for treatment or prevention of disorders
CC of, or surgical or cosmetic repair of, such muscle tissues, by
CC administering a hedgehog polypeptide or ptc therapeutic. The
CC disorder may be muscle atrophy, in particular skeletal muscle
CC atrophy or cardiac muscle atrophy, cachexia, or muscular myopathy

CC (all claimed). The hedgehog polypeptide or ptc therapeutic can
CC inhibit growth of myoblastic-derived tissue to provide treatment of
CC hyperplastic or neoplastic growth of muscle tissue such as in
CC myoblastic sarcoma (also claimed). The hedgehog therapeutic
CC preferably comprises at least a bioactive extracellular portion of
CC a hedgehog protein (see AAV05510-19) encoded by a vertebrate hedgehog
CC gene (see AAX25098-107), especially a human hedgehog gene.
XX
SQ Sequence 425 AA;
Query Match 100.0%; Score 2218; DB 20; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.7e-227;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 MYEMLLTRLLVGFICALLVSSGLTCGPGRGIGKRHPKRLPLAYKQIPVAEKTLG 60
Db 1 MYEMLLTRLLVGFICALLVSSGLTCGPGRGIGKRHPKRLPLAYKQIPVAEKTLG 60
Oy 61 ASGRYEGKITRNSERKELTPNYPDIIFKDEENTGADRLMTORCKDLNALAISVNW 120
Db 61 ASGRYEGKITRNSERKELTPNYPDIIFKDEENTGADRLMTORCKDLNALAISVNW 120
Oy 121 PGVKLRVTEGMDGDHSESLHTEGRAVDITTSDDRSKYGLARLAVAGFDWYYES 180
Db 121 PGVKLRVTEGMDGDHSESLHTEGRAVDITTSDDRSKYGLARLAVAGFDWYYES 180
Oy 181 KAHHCSVKAENSVAAKSGCFPGSATVHLEHGCTKLVKDLSFGDRVLAADADGRLLYSD 240
Db 181 KAHHCSVKAENSVAAKSGCFPGSATVHLEHGCTKLVKDLSFGDRVLAADADGRLLYSD 240
Oy 241 FLTFIDRMDSSRKLFFYIETROPARALLTAHLLFVAPDHNSGATGSGALFASNV 300
Db 241 FLTFIDRMDSSRKLFFYIETROPARALLTAHLLFVAPDHNSGATGSGALFASNV 300
Oy 301 KEGQRYVYGEQGQQLLPASVHSVSLREASGAYAPLTAOGTLLINRVLASCAVYEHS 360
Db 301 KEGQRYVYGEQGQQLLPASVHSVSLREASGAYAPLTAOGTLLINRVLASCAVYEHS 360
Oy 361 WAHMAFAPFRLAOGLLAALCPDGAIPTAATTTGIMHYSRLYRIGSWLDGALPLGM 420
Db 361 WAHMAFAPFRLAOGLLAALCPDGAIPTAATTTGIMHYSRLYRIGSWLDGALPLGM 420
Oy 421 VAPAS 425
Db 421 VAPAS 425
RESULT 4
AAW97765
ID AAW97765 standard; Protein; 425 AA.
AC AAW97765;
XX
XX 21-MAY-1999 (first entry)
DE Chicken Sonic hedgehog (Shh) protein.
XX
XX Sonic hedgehog; Shh protein; chicken; dopaminergic; GABA-ergic;
KW ptc therapeutic; patched; signal transduction; Parkinson's disease;
KW Huntington's disease; amyotrophic lateral sclerosis;
KW cerebral ischemia; hypoxia; neuroprotective; therapy.
OS Gallus domesticus.
XX
XX MO9904775-A2.
PN
XX 04-FEB-1999.
PD
XX 24-JUL-1998; 98WO-US15419.
PF
XX 24-JUL-1997; 97US-0900220.
PR
XX (ONTO-) ONTOGENY INC.
PA

XX Mahanthappa NK, Miao N, Pang K, Wang M;
PI WPI: 1999-142578/12.
XX N-PSDB: AAX07271.
DR
XX
XX Increasing the survival of neuronal, dopaminergic and GABA-nergic
PT cells - by using a ptc therapeutic such as a protein kinase
PT inhibitor, or an agent derived from hedgehog polypeptides, useful in
PT the treatment of Parkinson's disease
XX
PS Disclosure: Page 83-85; 138pp; English.
XX
XX This polypeptide is chicken Shh Sonic hedgehog protein. The
CC invention is based on the finding that hedgehog proteins are useful
CC as protective agents in the treatment and prophylaxis of
CC neurodegenerative disorders resulting from the loss of dopaminergic
CC and/or GABA-nergic neurons, or the general loss of tissue from the
CC substantia nigra. Exemplary disorders include Parkinson's disease,
CC Huntington's disease (both claimed), amyotrophic lateral sclerosis
CC and cerebral ischaemia. The invention relates to hedgehog
CC therapeutics (i.e. hedgehog polypeptides and gene therapy constructs
CC e.g. constructs encoding recombinant hedgehog polypeptides and
CC trans-activation constructs for altering hedgehog gene regulatory
CC sequences) and ptc therapeutics (i.e. agents which mimic the effect
CC of naturally occurring hedgehog proteins on patched signalling)
CC that are effective in both human and animal subjects. Human Ihh
CC and Dhh polypeptides (see AAW97763-64) are preferred. The products
CC can also be used for the maintenance of differentiated neurons in
CC cultures, and to enhance the implantation of such neuronal cells in
CC an animal. They can be used to prevent or treat neurodegenerative
CC conditions arising from the use of certain drugs, and in the
CC prevention and/or treatment of hypoxia, e.g. as a neuroprotective
CC agent.
XX
XX Sequence 425 AA;
SQ
Query Match 100.0%; Score 2218; DB 20; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.7e-227;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVEMLLRILLVGFICALLVSSGLTCGPGGIGRRRHPKLTPLAYKQFIPNVAEKLIG 60
DB 1 MVEMLLRILLVGFICALLVSSGLTCGPGGIGRRRHPKLTPLAYKQFIPNVAEKLIG 60
QY 61 ASGREGKITRNSERFKELTPNYNDIIFKDEENCGARLMTORCKDKLNALATISVMQW 120
DB 61 ASGREGKITRNSERFKELTPNYNDIIFKDEENCGARLMTORCKDKLNALATISVMQW 120
QY 121 PGVKLRVTEGWDGHHSESLHYEGRAVDITTSDDRSKYGMRLARLVEAGFDWVYES 180
DB 121 PGVKLRVTEGWDGHHSESLHYEGRAVDITTSDDRSKYGMRLARLVEAGFDWVYES 180
QY 181 KAHITCSVKAENSVAAKSGGCPGSATVLEHGCTKLVKDISPGGRVLAADAGLLYSD 240
DB 181 KAHITCSVKAENSVAAKSGGCPGSATVLEHGCTKLVKDISPGGRVLAADAGLLYSD 240
QY 241 FLTFIDRMDSSKRLFYVETROPARLLLTAAHLIFVAPHOHNSQATSTSGOALFASNV 300
DB 241 FLTFIDRMDSSKRLFYVETROPARLLLTAAHLIFVAPHOHNSQATSTSGOALFASNV 300
QY 301 KPGQAVYVLGEGGOOLLFASVHSVSLREASGAVAPLTAOCTILINRVLASCYAVIEHS 360
DB 301 KPGQAVYVLGEGGOOLLFASVHSVSLREASGAVAPLTAOCTILINRVLASCYAVIEHS 360
QY 361 WAHNAFAFRLAOGLLALCPDGAIPATAATTGTHWYSRLLRGSAVNDGDLAHPG 420
DB 361 WAHNAFAFRLAOGLLALCPDGAIPATAATTGTHWYSRLLRGSAVNDGDLAHPG 420
QY 421 VAPAS 425
DB 421 VAPAS 425

RESULT 5
AAW94468
ID AAW94468 standard; Protein; 425 AA.
XX
XX
AC AAW94468;
XX
DT 29-APR-1999 (first entry)
XX
DE Chicken Shh hedgehog protein sequence.
XX
KM Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;
KM brain infarction; cerebral infarction; transient ischaemic attack;
KM stroke; cerebral infarct volume; spinal cord; oedema; trauma;
KM haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.
XX
OS Gallus sp.
XX
XX WO9900117-A2.
XX
XX 07-JAN-1999.
XX
XX 26-JUN-1998; 98WO-US13387.
XX
XX 27-JUN-1997; 97US-0883656.
XX
PA (ONTO-) ONTOGENY INC.
XX
XX Mahanthappa NK;
XX
XX WPI: 1999-095458/08.
DR N-PSDB: AAX16182.
XX
PT Method for limiting damage to neurons caused by ischaemic or epoxic
PT conditions - is used for the treatment and prevention of e.g.
PT cerebral infarction, stroke and transient ischaemic attacks
XX
XX
PS Disclosure: Page 63-64; 104pp; English.
XX
XX A method has been developed for limiting the damage to neuronal cells by
CC ischaemic or epoxic conditions by administering a ptc (patched)
CC therapeutic agent to reduce cerebral infarct volume (CIV). Damage to
CC neuronal cells can also be limited by administering a gene activation
CC construct which recombines with the genomic hedgehog gene to provide a
CC heterologous transcription regulator linked to the coding region of this
CC gene. Administration of the ptc therapeutic agent is used to protect or
CC cerebral tissues against ischaemic injury; to treat cerebral infarct or
CC ischaemia, stroke (thrombotic or embolic) and transient ischaemic
CC attacks. It may also be used as a prophylactic in many other cases of
CC injury to the brain or spinal cord, oedema caused by trauma, haemorrhage
CC and encephalomyelitis, or in conjunction with (coronary bypass) surgery.
CC Treatment (which may be prophylactic) is used where ischaemic/epoxic
CC conditions may cause cerebral hypoxia, or progressive loss of neurons
CC due to oxygen depletion, including in patients with hypotension. The
CC treatment reduces CIV by at least 25, particularly at least 70%. The
CC present sequence represents a hedgehog sequence given in the present
CC invention.
XX
XX Sequence 425 AA;
SQ
Query Match 100.0%; Score 2218; DB 20; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.7e-227;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVEMLLRILLVGFICALLVSSGLTCGPGGIGRRRHPKLTPLAYKQFIPNVAEKLIG 60
DB 1 MVEMLLRILLVGFICALLVSSGLTCGPGGIGRRRHPKLTPLAYKQFIPNVAEKLIG 60
QY 61 ASGREGKITRNSERFKELTPNYNDIIFKDEENCGARLMTORCKDKLNALATISVMQW 120
DB 61 ASGREGKITRNSERFKELTPNYNDIIFKDEENCGARLMTORCKDKLNALATISVMQW 120
QY 121 PGVKLRVTEGWDGHHSESLHYEGRAVDITTSDDRSKYGMRLARLVEAGFDWVYES 180

|||||
Db 121 PGVKLRVTEGMEDGHSESLHYEGRAVDITTSDDRKYGMRLAFLAEGFDWYYES 180
Oy 181 KAHHCYKAENSVAKSGCCPGSATVHLEHGTRKLVNLSPGDRVLAADADGRLLYS 240
Db 181 KAHHCYKAENSVAKSGCCPGSATVHLEHGTRKLVNLSPGDRVLAADADGRLLYS 240
Oy 241 FLTFLDRMDSRKLFYIETROPRAALLTAAHLLEFVAPQHNOSEATGSGQALFASNV 300
Db 241 FLTFLDRMDSRKLFYIETROPRAALLTAAHLLEFVAPQHNOSEATGSGQALFASNV 300
Oy 301 KPGQRYVVLGEGGQQLLPASVHSVSLREASGAYAPLTAGTILLNRVLASCYAYIEHS 360
Db 301 KPGQRYVVLGEGGQQLLPASVHSVSLREASGAYAPLTAGTILLNRVLASCYAYIEHS 360
Oy 361 WAHMAFAPFRLAQGLLAALCPDGAIPTAATTGTHWYSRLYRIGSWLDDALHPLGM 420
Db 361 WAHMAFAPFRLAQGLLAALCPDGAIPTAATTGTHWYSRLYRIGSWLDDALHPLGM 420
Oy 421 VAPAS 425
Db 421 VAPAS 425

RESULT 6
AA95972
ID AA95972 standard; Protein; 425 AA.

XX AC AA95972;

DT 05-DEC-2000 (first entry)

DE Chicken Sonic hedgehog Shh protein.

XX KW Sonic hedgehog: Shh; chicken; agonist; antagonist;
KW lipid modulator; vacuole; cholesterol disorder; lipid disorder;
KW lipid metabolism; lipid storage; lipid transport; apolipoprotein;
KW triglyceride; hypercholesterolemia; abetalipoproteinemia;
KW hypobetalipoproteinemia; chylomicron retention; Anderson's disease;
KW fat absorption; atherosclerosis; obesity; weight loss;
KW vitamin A disorder; vitamin E disorder; antiilepemia; anorectic;
KW antidiarrhoeal; gene therapy; diagnosis.

XX OS Gallus sp.

XX PM MO200051628-A2.

XX PD 08-SEP-2000.

XX PF 03-MAR-2000; 2000WO-US05662.

XX PR 03-MAR-1999; 99US-0122640.

XX PR 15-MAR-1999; 99US-0124446.

XX PA (BIOU) BIOGEN INC.

XX PI Burkly L, Wang LC;

XX DR WPI: 2000-611340/58.

XX DR P-PSDB: AAA50446.

XX PT Use of lipid modulators (e.g. hedgehog agonists or antagonists) for
XX PT modulating lipid metabolism and storage, especially useful for treating
XX PT lipid metabolism or cholesterol disorders, e.g. obesity or
XX PT hypercholesterolemia -

XX PS Disclosure: Page 107-109; 136pp; English.

XX CC The present sequence of that of chicken Sonic hedgehog (Shh)
XX CC protein. The invention provides claimed methods for modulating
XX CC lipid metabolism, for modulating vacuole formation in intestinal
XX CC epithelial cells, for modulating the accumulation of fat in
XX CC intestinal epithelial cells, for treating a cholesterol disorder

CC and for treating a lipid metabolism disorder in an animal
CC (especially a human) by administering a lipid modulator selected
CC from a hedgehog antagonist or hedgehog agonist. In particular, the
CC lipid metabolism disorder is a lipid storage disorder, a lipid
CC transport disorder, an apolipoprotein disorder, a triglyceride
CC disorder, e.g. a triglyceride metabolism disorder, a triglyceride
CC transport disorder or a triglyceride storage disorder; a
CC diet-induced hypercholesterolemia, hypercholesterolemia,
CC chylomicron-retention disorder, hypobetalipoproteinemia; a
CC absorption disorder, e.g. obesity or associated with weight loss,
CC normotriglyceridemic abetalipoproteinemia, an apolipoprotein-B100
CC deficiency, a fat soluble vitamin disorder, where the fat soluble
CC vitamin is vitamin A or E, or atherosclerosis (all claimed). The
CC hedgehog antagonist binds to the hedgehog receptor, but does not
CC elicit a response. It is preferably a hedgehog mimetic, a modified
CC hedgehog protein, e.g. an inactive hedgehog variant, or a humanised
CC antibody. The methods are useful in preventing these disorders or
CC protecting a subject from these disorders. The hedgehog antagonist
CC and agonist are also useful in diagnosis and research associated
CC with these disorders. The lipid modulators may also be used as a
CC part of a gene therapy protocol to deliver polynucleotides encoding
CC these lipid modulators.

XX SQ Sequence 425 AA;

XX Query Match 100.0%; Score 2218; DB 21; Length 425;

XX Best Local Similarity 100.0%; Pred. No. 1.7e-227; Mismatches 0;

XX Matches 425; Conservative 0; Indels 0; Gaps 0;

Oy 1 WVEMLLRLILVGFICALVSSGLTGGGRIKRRHPPKLTPLAYKOFIPNVAETKLG 60

Db 1 WVEMLLRLILVGFICALVSSGLTGGGRIKRRHPPKLTPLAYKOFIPNVAETKLG 60

Oy 61 ASGRYEGKTRNSERKELTPVNPDIIFKDEENTGADRLMTORCKDKINALAISVMNQ 120

Db 61 ASGRYEGKTRNSERKELTPVNPDIIFKDEENTGADRLMTORCKDKINALAISVMNQ 120

Oy 121 PGVKLRVTEGMEDGHSESLHYEGRAVDITTSDDRKYGMRLAFLAEGFDWYYES 180

Db 121 PGVKLRVTEGMEDGHSESLHYEGRAVDITTSDDRKYGMRLAFLAEGFDWYYES 180

Oy 181 KAHHCYKAENSVAKSGCCPGSATVHLEHGTRKLVNLSPGDRVLAADADGRLLYS 240

Db 181 KAHHCYKAENSVAKSGCCPGSATVHLEHGTRKLVNLSPGDRVLAADADGRLLYS 240

Oy 241 FLTFLDRMDSRKLFYIETROPRAALLTAAHLLEFVAPQHNOSEATGSGQALFASNV 300

Db 241 FLTFLDRMDSRKLFYIETROPRAALLTAAHLLEFVAPQHNOSEATGSGQALFASNV 300

Oy 301 KPGQRYVVLGEGGQQLLPASVHSVSLREASGAYAPLTAGTILLNRVLASCYAYIEHS 360

Db 301 KPGQRYVVLGEGGQQLLPASVHSVSLREASGAYAPLTAGTILLNRVLASCYAYIEHS 360

Oy 361 WAHMAFAPFRLAQGLLAALCPDGAIPTAATTGTHWYSRLYRIGSWLDDALHPLGM 420

Db 361 WAHMAFAPFRLAQGLLAALCPDGAIPTAATTGTHWYSRLYRIGSWLDDALHPLGM 420

Oy 421 VAPAS 425

Db 421 VAPAS 425

XX RESULT 7

XX AA95281
XX ID AA95281 standard; Protein; 425 AA.

XX AC AA95281;

XX DT 12-SEP-2000 (first entry)

XX DE Chicken Sonic hedgehog Shh protein.

XX Sonic hedgehog; Shh; chicken; excitotoxicity; Parkinson's disease;
 KW Huntington's disease; neuronal degeneration; neuroprotective;
 KM dopaminergic; GABAergic; substantia nigra; therapy.
 XX
 OS Gallus sp.
 XX
 PN W0200035948-A1.
 XX
 PD 22-JUN-2000.
 XX
 PF 03-DEC-1999; 99WO-US28721.
 XX
 PR 03-DEC-1998; 98WO-US25676.
 PR 27-JAN-1999; 99US-0238243.
 PR 03-JUN-1999; 99US-0325602.
 XX
 PA (BIOJ) BIOGEN INC.
 PA (ONTO-) ONTOGENT INC.
 XX
 PI Galdes A, Mahanthappa N;
 XX
 DR WPI: 2000-431570/37.
 DR N-PSDB; AAA27876.
 XX
 PT Treating disorders involving exotoxicity, e.g. trauma, hypoglycemia,
 PT senile dementia and Korsakoff's disease, by using lipophilic modified
 PT hedgehog polypeptide -
 XX
 PS Disclosure: Page 149-151; 174pp; English.
 XX
 CC The present sequence of that of chicken Sonic hedgehog (Shh)
 CC protein. The invention relates to a method for promoting the
 CC survival and/or functional performance of neuronal cells,
 CC especially substantia nigra, dopaminergic or GABAergic neurons that
 CC are susceptible to exotoxicity, by contacting the cells, in vitro
 CC or in vivo, with a lipophilic (e.g. cholesterol) modified hedgehog
 CC polypeptide. The method is used to treat or prevent Parkinson's
 CC disease, Huntington's disease, domoic acid poisoning, spinal cord
 CC trauma, hypoglycemia, mechanical trauma to the nervous system,
 CC senile dementia, Korsakoff's disease, schizophrenia, AIDS dementia,
 CC multi-infarct dementia, mood disorders, depression, chemical
 CC toxicity, neuronal damage associated with uncontrolled seizures
 CC such as epileptic seizures, neuronal injury associated with HIV and
 CC AIDS, neurodegeneration associated with Down's syndrome,
 CC neuropathic pain syndrome, olivopontocerebral atrophy, amyotrophic
 CC lateral sclerosis, mitochondrial abnormalities, Alzheimer's disease,
 CC hepatic encephalopathy, Tourette's syndrome and drug addiction (all
 CC claimed). The lipophilic modified hedgehog polypeptide is also
 CC useful for promoting survival and/or functional performance of
 CC neuronal cells susceptible to exotoxicity.
 CC
 XX
 SQ Sequence 425 AA:
 Query Match 100.0%; Score 2218; DB 21; Length 425;
 Best Local Similarity 100.0%; Pred. No. 1.7e-227;
 Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 241 LFLFLDRMSSRKLFYIETROPARLLTPAHLFLFAPQHNQSEAGTSGQALFASNV 300
 |||||||
 DB 241 FLFLDRMSSRKLFYIETROPARLLTPAHLFLFAPQHNQSEAGTSGQALFASNV 300
 OY 301 KPGGRVYVLGGGQQLPASHVSVSLREASGAYAPLTAQCTIILINRYLASCYAVIEHS 360
 |||||||
 DB 301 KPGGRVYVLGGGQQLPASHVSVSLREASGAYAPLTAQCTIILINRYLASCYAVIEHS 360
 OY 361 WAHNAFAPFRLAOGLLAALCPDGAIPTAATTTGTGHHVSRLLYRIGSVLDPDGLHPHGM 420
 |||||||
 DB 361 WAHNAFAPFRLAOGLLAALCPDGAIPTAATTTGTGHHVSRLLYRIGSVLDPDGLHPHGM 420
 OY 421 VAPAS 425
 |||||
 DB 421 VAPAS 425
 RESULT 8
 ID AAY96243
 ID AAY96243 standard; Protein: 425 AA.
 XX
 AC AAY96243;
 XX
 DT 11-SEP-2000 (first entry)
 XX
 DE Partial chicken Shh.
 XX
 KW Chicken; sonic hedgehog; Shh; neuromuscular disorder; neuropathy;
 KW Gullain-Barre syndrome; peripheral neuropathy; diabetes; alcoholism;
 KW chronic inflammatory demyelinating polyneuropathy; CIDP;
 KW gene therapy; infection; inflammation; hereditary neuropathy;
 KW Charcot-Marie-Tooth disease; vasculitis; lung cancer; tumour;
 KW multiple myeloma; nutritional imbalance; kidney disease;
 KW hypothyroid neuropathy; trauma; Refsum's disease; Abetalipoproteinemia;
 KW Fabry's disease; Krabbe's disease; Metachromatic leukodystrophy;
 KW Amyotrophic lateral sclerosis; ALS; Miller-Fisher syndrome; amyloidosis;
 KW hereditary sensory neuropathy Type II; HSN II; B-cell lymphoma;
 KW Waldenstrom's Macroglobulinemia; Chronic lymphocytic leukaemia;
 KW neuroprotective; cytoprotective; patched-mediated signal transduction.
 XX
 OS Gallus domesticus.
 XX
 PN W0200027422-A2.
 XX
 PD 18-MAY-2000.
 XX
 PF 08-NOV-1999; 99WO-US26334.
 XX
 PR 06-NOV-1998; 98US-0187387.
 XX
 PA (BIOJ) BIOGEN INC.
 PA (ONTO-) ONTOGENT INC.
 XX
 PI Galdes A, Mahanthappa N;
 XX
 DR WPI: 2000-387341/33.
 DR N-PSDB; AAA30274.
 XX
 PT Novel method of preventing deterioration of peripheral nerves, useful
 PT for treating or preventing neuropathy, e.g. where associated with
 PT diabetes or viral infection, by administering hedgehog or patched agent
 PT -
 PS Claim 7; Page 127-128; 152pp; English.
 XX
 CC The present sequence is the partial chicken sonic hedgehog protein, Shh.
 CC This sequence inhibits expression of the patched gene which has been
 CC implicated in neuromuscular disorders (neuropathies). This sequence may
 CC therefore be used for treating neuromuscular disorders i.e. preventing
 CC degradation in function of motor or sensory nerves and protecting
 CC peripheral nerve cells under conditions that normally cause neuropathy.


```

Db 361 WAHMAFAFPRLAAGLLAALCPDGAIPATAITTTGTHWYSLRLYRIGSVNLDDADLHPGLM 420
QY 421 VAPAS 425
Db 421 VAPAS 425

RESULT 10
AAG65743
ID AAG65743 standard; Protein: 425 AA.
XX
AC AAG65743:
XX
DT 07-JAN-2002 (first entry)
XX
DE Chicken sonic hedgehog (Shh) polypeptide.
XX
KW Adipocyte; hedgehog polypeptide; desert hedgehog; indian hedgehog; Dh;
KW Ihn; sonic hedgehog; Shh; therapeutic; cytostatic; tumor.
XX
OS Gallus sp.
XX
PN MO200164238-A2.
XX
PD 07-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US06450.
XX
PR 29-FEB-2000; 2000US-186058P.
XX
PA (CURT-) CURIS INC.
XX
PI Zehentner B, Leser-Reiff U, Bartscher H;
XX
DR WPI: 2001-607352/69.
XX
DR N-PSDB; AA166771.
XX
PT Method for regulating formation and/or maintenance of adipocyte tissue
PT by contacting pre-adipocyte or adipocyte cells with a hedgehog
PT polypeptide or plc therapeutic
XX
PS Disclosure; Page 90-92; 132pp; English.
XX
CC The invention provides a method for regulating formation and/or
CC maintenance of adipocyte tissue that comprises contacting pre adipocyte
CC or adipocyte cells with a hedgehog polypeptide or plc therapeutic. The
CC method is used for regulating the growth state of an adipocyte stem/
CC progenitor cell, and treating or preventing disorders of, or surgical or
CC cosmetic repair of, adipocyte tissues, e.g. for treating or preventing
CC hyperplastic or neoplastic conditions affecting adipocyte tissue, such
CC as soft tissue tumors, especially adipose cell tumors, e.g. lipomas,
CC fibrolipomas, lipoblastomas, lipomatosis, fibromas, hemangiomas and/or
CC liposarcomas. Hedgehog polypeptides can be used in combination with other
CC therapeutic agents. The present sequence represents a chicken sonic
CC hedgehog (Shh) polypeptide.
XX
SQ Sequence 425 AA:
Query Match 100.0%; Score 2218; DB 22; Length 425;
Best Local Similarity 100.0%; Pred. No. 1,7e-227;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVEMLLRLILVGFICALVSSGLTCGPGNGICGRHRHKKLTPLAYKQIFPNAEKTIG 60
Db 1 MVEMLLRLILVGFICALVSSGLTCGPGNGICGRHRHKKLTPLAYKQIFPNAEKTIG 60
QY 61 ASGREKGTNSERFKELTPNYNDLIFKDEENTGADRLMTQKDKLNLALISVNMQW 120
Db 61 ASGREKGTNSERFKELTPNYNDLIFKDEENTGADRLMTQKDKLNLALISVNMQW 120
QY 121 PGVKLRVEGMDGDGHHSEESLHYEGRAVDITTSRDRSKTGMLARLAVEAGFDWVYES 180

```

```

Db 121 PGVKLRVEGMDGDGHHSEESLHYEGRAVDITTSRDRSKTGMLARLAVEAGFDWVYES 180
QY 181 KAHHC SVKAENSVAAKSGCGPGSATVHLEHGTRKLVKDSLPGDRLVAADAGRLYSD 240
Db 181 KAHHC SVKAENSVAAKSGCGPGSATVHLEHGTRKLVKDSLPGDRLVAADAGRLYSD 240
QY 241 FLTFIDRRDSSKRLFYVETROPARLLITAAHLFVAPQHNQSEATSTSQALFASNV 300
Db 241 FLTFIDRRDSSKRLFYVETROPARLLITAAHLFVAPQHNQSEATSTSQALFASNV 300
QY 301 KPGQRYVYLGEQGQULLPASVHVSILREASGAVAPLTAOGTILNRYLASCYAVTEHS 360
Db 301 KPGQRYVYLGEQGQULLPASVHVSILREASGAVAPLTAOGTILNRYLASCYAVTEHS 360
QY 361 WAHMAFAFPRLAAGLLAALCPDGAIPATAITTTGTHWYSLRLYRIGSVNLDDADLHPGLM 420
Db 361 WAHMAFAFPRLAAGLLAALCPDGAIPATAITTTGTHWYSLRLYRIGSVNLDDADLHPGLM 420
QY 421 VAPAS 425
Db 421 VAPAS 425

RESULT 11
AAB85733
ID AAB85733 standard; Protein: 425 AA.
XX
AC AAB85733:
XX
DT 29-OCT-2001 (first entry)
XX
DE Chicken sonic hedgehog (Shh) polypeptide.
XX
KW Hedgehog protein; sonic hedgehog; Shh; indian hedgehog; Ihn; Dh;
KW desert hedgehog; cell differentiation; chicken.
XX
OS Gallus sp.
XX
PN US6271363-B1.
XX
PD 07-AUG-2001.
XX
PF 20-OCT-1997; 97US-0954698.
XX
PR 05-JUN-1995; 95US-0462386.
PR 30-DEC-1993; 93US-0176427.
PR 14-DEC-1994; 94US-0356060.
PR 04-MAY-1995; 95US-0435093.
XX
PA (HARD ) HARVARD COLLEGE.
PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.
XX
PI Ingham PW, McMahon AP, Tablin CJ;
XX
DR WPI: 2001-456723/49.
XX
DR N-PSDB; AAH76107.
XX
PT Novel nucleic acid encoding a hedgehog polypeptide, used to produce the
PT polypeptide, which is used to promote proliferation, survival, and/or
PT differentiation of neuronal and mesodermal tissue -
XX
PS Claim 1; Column 117-120; 118pp; English.
XX
CC The invention relates to nucleic acids encoding hedgehog proteins
CC selected from sonic hedgehog (Shh), indian hedgehog (Ihn), desert
CC hedgehog (Dhh) polypeptides. The hedgehog genes are involved in the
CC formation of ordered spatial arrangements of differentiated tissue in
CC vertebrates. The nucleic acid sequences are useful for producing hedgehog
CC proteins, used for promoting differentiation of, or survival of
CC differentiated, neuronal cells, and for promoting proliferation, survival
CC or differentiation of mesenchymal, endodermal or ectodermal tissue,
CC particularly chondrocytes, or testicular germ line cells. The present
CC sequence represents a chicken Shh polypeptide.

```

[illegible]

PT	Modulating immune function comprises administration of a hedgehog or	
PT	ptc agent, for treating e.g. diabetes, eczematous dermatitis, urticaria	
PT	or vasculitis .	
XX		
PS	Claim 4; Page 77-78; 105pp; English.	
XX		
CC	The present sequence represents a hedgehog (Shh) polypeptide. Hedgehog	
CC	gene products and signal transduction pathways involving hedgehog are	
CC	involved in the maturation of T lymphocytes. The specification describes	
CC	a method for modulating immune function, by administration of a hedgehog	
CC	or patched (ptc) polypeptide, agonists or antagonists. The method is	
CC	used to treat disorders affecting the regulation of lymphocytes,	
CC	particularly maturation and/or activation of T lymphocytes. It is used	
CC	to treat bacterial or viral infection, diabetes, nutritional	
CC	deficiencies, graft rejection or other hyperacute response such as	
CC	kidney, heart, lung, bone marrow spleen skin or cornea transplant or	
CC	autoimmune disorders such as multiple sclerosis, psoriasis or atopic	
CC	dermatitis. The method is used to treat inflammatory, proliferative and	
CC	hyperproliferative diseases, as well as cutaneous manifestations of	
CC	immunological disorders such as eczematous dermatitis, urticaria,	
CC	vasculitis and scleroderma.	
XX		
XX	Sequence 425 AA;	
XX		
Query Match	100.0%; Score 2218; DB 22; Length 425;	
Best local similarity	100.0%; Pred. No. 1.7e-227;	
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 MVEMLLRLILVGFICALLVSSGLTGPCGPGICGRRRPKLTPLAYKOFIPNVAEKTG	60
Db	1 MVEMLLRLILVGFICALLVSSGLTGPCGPGICGRRRPKLTPLAYKOFIPNVAEKTG	60
QY	61 ASGREGKITTENSEBFKLTNPVNDIIFKPEENTGADRLMTORCKDKLNALATISVMNW	120
Db	61 ASGREGKITTENSEBFKLTNPVNDIIFKPEENTGADRLMTORCKDKLNALATISVMNW	120
QY	121 PGVKLRVTEGDEDEHSHSESLHYEGRAVDITSDRDSKYGMLRLAVEAGFDWYYES	180
Db	121 PGVKLRVTEGDEDEHSHSESLHYEGRAVDITSDRDSKYGMLRLAVEAGFDWYYES	180
QY	121 PGVKLRVTEGDEDEHSHSESLHYEGRAVDITSDRDSKYGMLRLAVEAGFDWYYES	180
Db	121 PGVKLRVTEGDEDEHSHSESLHYEGRAVDITSDRDSKYGMLRLAVEAGFDWYYES	180
QY	181 KAHITCSYKAENSVAAKSGGCPGSAVYHLEHGTRKLYKDSLPGGRVLAADADGRLYSD	240
Db	181 KAHITCSYKAENSVAAKSGGCPGSAVYHLEHGTRKLYKDSLPGGRVLAADADGRLYSD	240
QY	241 FLTFLDRDSSRKLFYVETROPARRLTLTAHLFLFVPOHNSPATSGOALFASNV	300
Db	241 FLTFLDRDSSRKLFYVETROPARRLTLTAHLFLFVPOHNSPATSGOALFASNV	300
QY	301 KPGGRVYVLEGGOQLLPASVHSVLSREESAGAVPLTAOGTILINRVLASCYAVIEHS	360
Db	301 KPGGRVYVLEGGOQLLPASVHSVLSREESAGAVPLTAOGTILINRVLASCYAVIEHS	360
QY	361 WAHMAFAFRLAOGLLALCPDGAIPATAATTTGTHWYSLRLYRGSWLDDAHPGLM	420
Db	361 WAHMAFAFRLAOGLLALCPDGAIPATAATTTGTHWYSLRLYRGSWLDDAHPGLM	420
QY	421 VAPAS 425	
Db	421 VAPAS 425	
RESULT 13		
ID	AAE04682	
AC	AAE04682 standard; Protein: 425 AA.	
XX		
XX	AAE04682;	
XX		
XX	04-SEP-2001 (first entry)	
XX		
XX	Chicken sonic hedgehog (Shh) protein.	
XX		
XX	Chicken: hedgehog protein; nootropic; neuroprotective; anticonvulsant;	

KW	Cytostatic; therapy; Alzheimer's disease; Parkinson's disease; injury;
KW	Huntington's chorea; amyotrophic lateral sclerosis; multiple sclerosis;
KM	neurous system aging; neurodegenerative disease; immunological disease;
KW	malignant glioma; medulloblastoma; neuroectodermal tumour; cancer;
KW	extracellular signalling protein.
XX	
OS	Gallus sp.
PN	WO200134654-A1.
PD	17-MAY-2001.
PP	
PE	02-NOV-2000; 2000WO-US30405.
PR	05-NOV-1999; 99US-0164025.
PA	(BIOJ) BIOGEN INC.
PI	Strauch K:
DR	WPI: 2001-329075/34.
DR	N-PSDB; AAD09029.
PT	Novel isolated hedgehog fusion polypeptide useful for treating
PT	neurological conditions such as Alzheimer's disease, Parkinson's
PT	disease, Huntington's chorea, amyotrophic lateral sclerosis, and
PT	multiple sclerosis -
PS	Disclosure; Page 111-112; 178pp; English.
XX	
CC	The present invention relates to hedgehog fusion proteins. Hedgehog
CC	proteins are a family of extracellular signaling proteins that regulate
CC	various aspects of embryonic development both in vertebrates and in
CC	invertebrates. Hedgehog fusion protein is useful for the prophylaxis or
CC	treatment of any condition or disease state for which a hedgehog or
CC	patched protein constituent is efficacious and in the diagnosis of
CC	constituents or conditions of disease states in biological system or
CC	specimens and for diagnostic purposes in non-physiological systems.
CC	Hedgehog fusion protein is useful for treating neurological conditions
CC	due to injury, aging of nervous system, including Alzheimer's disease,
CC	chronic neurodegenerative diseases of the nervous system, including
CC	Parkinson's disease, Huntington's chorea, amyotrophic lateral sclerosis
CC	and chronic immunological diseases of nervous system including multiple
CC	sclerosis and malignant gliomas, medulloblastomas, neuroectodermal
CC	tumours and to specifically target medical therapies against cancers and
CC	tumours which express the receptor for the protein. The present sequence
CC	is chicken sonic hedgehog (Shh) protein.
XX	
SQ	Sequence 425 AA:
	Query Match 100.0%; Score 2218; DB 22; Length 425;
	Best Local Similarity 100.0%; Pred. No. 1,7e+227;
	Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 WEMLLTRLILLVGITCALVSSGLTCGPGRGIGKRRHPKLTPLAYKOFIPNVAEKTIG 60
Db	1 WEMELLTRILLVGITICALLVSSGLTCGPGRGIGKRRHPKLTPLAYKOFIPNVAEKTIG 60
OY	61 ASGRGEKITRSEFKELTPYNNDDIFKDEENFGADLMORCKDKINALAISMMNQ 120
Db	61 ASGRGEKITRSEFKELTPYNNDDIFKDEENFGADLMORCKDKINALAISMMNQ 120
OY	121 PGVKLRVEGMDEDHHSSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWYYES 180
Db	121 PGVKLRVEGMDEDHHSSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWYYES 180
OY	181 KAHHCYSKAENSVAAKSGGCFPGSAGTVAHLHGRTKLYKDLSPGDGVILAADAGRLLYS 240
Db	181 KAHHCYSKAENSVAAKSGGCFPGSAGTVAHLHGRTKLYKDLSPGDGVILAADAGRLLYS 240
OY	241 FLTLFDRMDSSKKLEVVETROPARRLLITAAHLFLVAPQHNSQEAETGSTSQALFASNV 300
Db	241 FLTLFDRMDSSKKLEVVETROPARRLLITAAHLFLVAPQHNSQEAETGSTSQALFASNV 300

QY	301	KPGQRVYVGEQGOQLIPASVHSVSLREESGAYAPLTACTGILLINNVLASCAVIEEHS	360		
Db	301	KPGQRVYVGEQGOQLIPASVHSVSLREESGAYAPLTACTGILLINNVLASCAVIEEHS	360		
QY	361	MAWMAAPFPLAAGLLAALCPDGAIPFAATTTTGIIHWSRLRLRIGSWVLDGALHPFGM	420		
Db	361	MAWMAAPFPLAAGLLAALCPDGAIPFAATTTTGIIHWSRLRLRIGSWVLDGALHPFGM	420		
QY	421	VAPAS 425			
Db	421	VAPAS 425			
RESULT 14					
AAE05372	AAE05372 standard; Protein; 425 AA.				
XX	AAE05372;				
XX	12-SEP-2001 (first entry)				
DE	Chicken Sonic hedgehog protein.				
KW	Chicken: Sonic hedgehog: Shh; morphogenic signal; neuron; embryonic patterning; cell culture; cell differentiation; ischaemia; cell proliferative disorder; Intracerebral grafting; Huntington's chorea; neurological disorder; Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis; ALS; multiple sclerosis.				
OS	Gallus sp.				
FI	Key	Location/Qualifiers			
FT	Peptide	1..26			
FT	Protein	/label= signal_peptide 27..425			
FT		/label= Mature_Sh_h_protein			
PN	US6261786-B1.				
PD	17-JUL-2001.				
PF	02-JUL-1996;	96US-0674509.			
PR	30-DEC-1993;	93US-0176427.			
PR	14-DEC-1994;	94US-0356060.			
PR	04-MAY-1995;	95US-0435093.			
PR	05-JUN-1995;	95US-0460900.			
PR	05-JUN-1995;	95US-0462386.			
PA	(IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.				
PA	(HARD) HARVARD COLLEGE.				
PI	Mariño V, Tabin CJ, Ingham PW, McMahon AP;				
DR	WPI: 2001-440859/47.				
DR	N-PSDB; AAD10146.				
XX	Screening compounds that potentiate or inhibit binding of hedgehog polypeptide to naturally occurring patched receptor, comprises				
PT	contacting polypeptide with receptor and test compound, and detecting change in binding -				
PT	Claim 2: Fig 5A: 127pp; English.				
PS	The present invention relates to assay for screening compounds that potentiate or inhibit binding of hedgehog polypeptide to naturally occurring patched receptor. The hedgehog proteins comprise morphogenic signals produced by embryonic patterning centres, and are involved in the formation and maintenance of ordered spatial arrangements of differentiated tissues in vertebrates, both adult and embryonic. The proteins can be used to generate and/or maintain an array of different vertebrate tissues both in vitro and in vivo. The invention also relates				

CC to a method for modulating growth, differentiation or survival of a
 CC mammalian cell (e.g. neuron, testicular cell) responsive to hedgehog
 CC induction. Hedgehog agonists and antagonists can be used in cell culture
 CC techniques to enhance survival and maintenance of neurons and various
 CC vertebrate organogenic pathways. The hedgehog gene is useful in
 CC determining whether a patient is at the risk of disorder characterised by
 CC unwanted cell proliferation or aberrant control of differentiation. The
 CC hedgehog proteins or mimetics can be used to induce foetal neurons
 CC especially neuronal stem cells in intracerebral grafting. The protein
 CC or its mimetic can be used in the treatment of neurological conditions
 CC e.g. injury to nervous system, ischaemia resulting from stroke,
 CC Alzheimer's disease, Parkinson's disease, Huntington's chorea,
 CC amyotrophic lateral sclerosis (ALS) and multiple sclerosis. The present
 CC sequence is chicken Sonic hedgehog (Shh) protein.

SQ Sequence 425 AA:

Query Match 100.0%; Score 2218; DB 22; Length 425;

Best Local Similarity 100.0%; Pred. No. 1.7e-227;

Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MVEMLLTRILLVGFICALLVSSGLTCGPGRGIGKRHPKRLPLAKOPIPNVAEXTLG 60
 DB 1 MVEMLLTRILLVGFICALLVSSGLTCGPGRGIGKRHPKRLPLAKOPIPNVAEXTLG 60
 OY 61 ASGRYEGKITRNSERFKEKLPNNYNDIIFKDEENTGADRLMTORCKDLNALAISVNNQW 120
 DB 61 ASGRYEGKITRNSERFKEKLPNNYNDIIFKDEENTGADRLMTORCKDLNALAISVNNQW 120
 OY 121 PGVKLRVTEGWDGHHSESLHYEGRAVDITTSDDRKSRYGMLARLAVAGFDWYYES 180
 DB 121 PGVKLRVTEGWDGHHSESLHYEGRAVDITTSDDRKSRYGMLARLAVAGFDWYYES 180
 OY 181 KAHHGSVKAENSVAAKSGGCPGSAIVHLEHGRTKLVKLSFGDRVLADADGRLLYSD 240
 DB 181 KAHHGSVKAENSVAAKSGGCPGSAIVHLEHGRTKLVKLSFGDRVLADADGRLLYSD 240
 OY 241 FLTFLDMDSSRKLFFYIETROPARALLTAHLFLVAPQHNSEATGSGALFASNV 300
 DB 241 FLTFLDMDSSRKLFFYIETROPARALLTAHLFLVAPQHNSEATGSGALFASNV 300
 OY 301 KPGORVYVLDGEGQOQLPASHVSLSREASGAYAPLTAOGTLLINRVLASCAVIEHS 360
 DB 301 KPGORVYVLDGEGQOQLPASHVSLSREASGAYAPLTAOGTLLINRVLASCAVIEHS 360
 OY 361 MAHMAFAPFRLAAGLLAALCPDAIPTAATTGIIHWYSRLYRIGSWVLDGDLHPFLGM 420
 DB 361 MAHMAFAPFRLAAGLLAALCPDAIPTAATTGIIHWYSRLYRIGSWVLDGDLHPFLGM 420
 OY 421 VAPAS 425
 DB 421 VAPAS 425

RESULT 15

AAB31217 AAB31217 standard; Protein: 425 AA.

AC AAB31217;

DT 20-APR-2001 (first entry)

XX Amino acid sequence of chicken sonic hedgehog protein (Shh).

XX Hedgehog related-protein; sonic hedgehog protein; Shh; Ischemia; stroke;

XX desert hedgehog protein; Dhh; Indian hedgehog protein; Ihh; neuron;

XX neurological condition; nervous system injury; tumour-induced injury;

XX aging; Alzheimer's disease; chronic neurodegenerative disease;

XX Parkinson's disease; Huntington's chorea; amyotrophic lateral sclerosis;

XX spinocerebellar degeneration; chronic immunological disease;

XX multiple sclerosis;

XX US6165747-A.
 PN 26-DEC-2000.
 XX
 PD 05-JUN-1995; 95US-0460900.
 XX
 PR 30-DEC-1993; 93US-0176427.
 PR 14-DEC-1994; 94US-0356060.
 PR 04-MAY-1995; 95US-0435093.
 XX
 PA (HARD) HARVARD COLLEGE.
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
 XX
 PI Ingham PM, McMahon AP, Tabin CJ, Marti-gorostiza E, Bumbrot DA;
 DR N-PSDB; AAC87074.
 XX
 PT Polynucleotides encoding hedgehog proteins, useful for treating
 PT diseases of nervous system such as Alzheimer's disease, Parkinson's
 PT disease, Huntington's chorea, amyotrophic lateral sclerosis, multiple
 PT sclerosis -
 XX
 PS Claim 10; Columns 135-138; 119pp; English.
 XX
 CC The present sequence represents a hedgehog related-protein. The
 CC specification describes a sonic hedgehog protein (Shh), a desert
 CC hedgehog protein (Dhh), and an Indian hedgehog protein (Ihh). The
 CC hedgehog polynucleotides are useful in diagnostic, in antisense
 CC therapy and in therapeutic assays for detecting and treating disorders
 CC involving, e.g., aberrant expression of vertebrate hedgehog homologue.
 CC Hedgehog polypeptides are useful therapeutically to enhance survival
 CC of neurons and other neuron cells and in treating neurological
 CC conditions deriving from acute, subacute, or chronic injury to the
 CC nervous system, including traumatic injury, chemical injury, vessel
 CC injury and deficits (such as the ischemia resulting from stroke),
 CC together with infectious/inflammatory and induced-injury, aging
 CC of the nervous system including Alzheimer's disease, chronic
 CC neurodegenerative diseases of the nervous system, including Parkinson's
 CC disease, Huntington's chorea, amyotrophic lateral sclerosis,
 CC spinocerebellar degenerations, and chronic immunological diseases of
 CC the nervous system or affecting the nervous system, including multiple
 CC sclerosis.
 CC
 XX
 SQ Sequence 425 AA:

Query Match 100.0%; Score 2218; DB 22; Length 425;

Best Local Similarity 100.0%; Pred. No. 1.7e-227;

Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MVEMLLTRILLVGFICALLVSSGLTCGPGRGIGKRHPKRLPLAKOPIPNVAEXTLG 60
 DB 1 MVEMLLTRILLVGFICALLVSSGLTCGPGRGIGKRHPKRLPLAKOPIPNVAEXTLG 60
 OY 61 ASGRYEGKITRNSERFKEKLPNNYNDIIFKDEENTGADRLMTORCKDLNALAISVNNQW 120
 DB 61 ASGRYEGKITRNSERFKEKLPNNYNDIIFKDEENTGADRLMTORCKDLNALAISVNNQW 120
 OY 121 PGVKLRVTEGWDGHHSESLHYEGRAVDITTSDDRKSRYGMLARLAVAGFDWYYES 180
 DB 121 PGVKLRVTEGWDGHHSESLHYEGRAVDITTSDDRKSRYGMLARLAVAGFDWYYES 180
 OY 181 KAHHGSVKAENSVAAKSGGCPGSAIVHLEHGRTKLVKLSFGDRVLADADGRLLYSD 240
 DB 181 KAHHGSVKAENSVAAKSGGCPGSAIVHLEHGRTKLVKLSFGDRVLADADGRLLYSD 240
 OY 241 FLTFLDMDSSRKLFFYIETROPARALLTAHLFLVAPQHNSEATGSGALFASNV 300
 DB 241 FLTFLDMDSSRKLFFYIETROPARALLTAHLFLVAPQHNSEATGSGALFASNV 300
 OY 301 KPGORVYVLDGEGQOQLPASHVSLSREASGAYAPLTAOGTLLINRVLASCAVIEHS 360
 DB 301 KPGORVYVLDGEGQOQLPASHVSLSREASGAYAPLTAOGTLLINRVLASCAVIEHS 360

Db 301 KPGORVYVLGEGGQULPASPVSILREBASCAYAPLTAQGTILINRVLASCYAVIEHS 360
Oy 361 WAHMAFAPRLAQGLLALCPDGAIPATAITTTGIHWYSRLIYRIGSWVLDGDAHP LGM 420
Db 361 WAHMAFAPRLAQGLLALCPDGAIPATAITTTGIHWYSRLIYRIGSWVLDGDAHP LGM 420
Oy 421 VAPAS 425
Db 421 VAPAS 425

Search completed: February 20, 2003, 10:10:17
Job time : 32.834 secs


```

Oy 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNLAISVMNQ 120
    |||||||
Db 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNLAISVMNQ 120
Oy 121 PGVKLRVTEGMEDEGHSEESLHYEGRAVDITTSRDRSKYGMRLARLAVEGFDMVYYES 180
    |||||||
Db 121 PGVKLRVTEGMEDEGHSEESLHYEGRAVDITTSRDRSKYGMRLARLAVEGFDMVYYES 180
Oy 181 KAHHCYSKAENSVAAKSGGCEPGSATVHLEHGTRKLVKDLSPGDRVLAADADGRLYSD 240
    |||||||
Db 181 KAHHCYSKAENSVAAKSGGCEPGSATVHLEHGTRKLVKDLSPGDRVLAADADGRLYSD 240
Oy 241 FLTFIDRMDSSRKLFYVETROPARLLITAAHLFVAPQHNQSEATGSGQALFASNV 300
    |||||||
Db 241 FLTFIDRMDSSRKLFYVETROPARLLITAAHLFVAPQHNQSEATGSGQALFASNV 300
Oy 301 KPGQRYVYLGEGGQQLLPASVHSVSLREASGAVAPLTAOGTILINRVLASCYAVIEHS 360
    |||||||
Db 301 KPGQRYVYLGEGGQQLLPASVHSVSLREASGAVAPLTAOGTILINRVLASCYAVIEHS 360
Oy 361 WAHMAFAFRLAOGLLALCPDGAIPTAATTTTGIHWYSRLLYRIGSVWLDGDALHPJGM 420
    |||||||
Db 361 WAHMAFAFRLAOGLLALCPDGAIPTAATTTTGIHWYSRLLYRIGSVWLDGDALHPJGM 420
Oy 421 VAPAS 425
    |||||
Db 421 VAPAS 425

```

RESULT 2

US-08-356-060A-8
Sequence (8) Application US/08356060A
Patent No. 5844079

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tabin, Clifford J.
TITLE OF INVENTION: Vertebrate and Uses Related Thereto
TITLE OF INVENTION: Vertebrate and Uses Related Thereto
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356, 060A
FILING DATE: 14-DEC-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-060A-8

Query Match 100.0%; Score 2218; DB 2; Length 425;
Best Local Similarity 100.0%; Pred. No. 1,5e-241;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 MEKMLLRILLVGFICALYSSGLTCGPGRGIGRRRPPKILTPLAYQOIFPNVAEKTG 60
    |||||||
Db 1 MEKMLLRILLVGFICALYSSGLTCGPGRGIGRRRPPKILTPLAYQOIFPNVAEKTG 60
Oy 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNLAISVMNQ 120
    |||||||
Db 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNLAISVMNQ 120
Oy 121 PGVKLRVTEGMEDEGHSEESLHYEGRAVDITTSRDRSKYGMRLARLAVEGFDMVYYES 180
    |||||||
Db 121 PGVKLRVTEGMEDEGHSEESLHYEGRAVDITTSRDRSKYGMRLARLAVEGFDMVYYES 180
Oy 181 KAHHCYSKAENSVAAKSGGCEPGSATVHLEHGTRKLVKDLSPGDRVLAADADGRLYSD 240
    |||||||
Db 181 KAHHCYSKAENSVAAKSGGCEPGSATVHLEHGTRKLVKDLSPGDRVLAADADGRLYSD 240
Oy 241 FLTFIDRMDSSRKLFYVETROPARLLITAAHLFVAPQHNQSEATGSGQALFASNV 300
    |||||||
Db 241 FLTFIDRMDSSRKLFYVETROPARLLITAAHLFVAPQHNQSEATGSGQALFASNV 300
Oy 301 KPGQRYVYLGEGGQQLLPASVHSVSLREASGAVAPLTAOGTILINRVLASCYAVIEHS 360
    |||||||
Db 301 KPGQRYVYLGEGGQQLLPASVHSVSLREASGAVAPLTAOGTILINRVLASCYAVIEHS 360
Oy 361 WAHMAFAFRLAOGLLALCPDGAIPTAATTTTGIHWYSRLLYRIGSVWLDGDALHPJGM 420
    |||||||
Db 361 WAHMAFAFRLAOGLLALCPDGAIPTAATTTTGIHWYSRLLYRIGSVWLDGDALHPJGM 420
Oy 421 VAPAS 425
    |||||
Db 421 VAPAS 425

```

RESULT 3

US-08-460-900C-8
Sequence 8, Application US/08460900C
Patent No. 6165747

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tabin, Clifford J.
APPLICANT: Bumcrot, David A.
APPLICANT: Marti-Gorostiza, Elisa
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HONG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460, 900C
FILING DATE: 5-JUNE-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/435,093
FILING DATE: 4-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/356,060
FILING DATE: 14-DEC-1994
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMV-006.05
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-900C-8

Query Match 100.0%; Score 2218; DB 4; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.5e-241;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVEMLLTLLILVGFICALLVSSGLTGCSPGRGIGKRHRPKKLPPLAYKOFIPNVAEKTIG 60
DB 1 MVEMLLTLLILVGFICALLVSSGLTGCSPGRGIGKRHRPKKLPPLAYKOFIPNVAEKTIG 60
QY 61 ASGRYEGKITRNSERFELTPNPNPDIIFKDEENTGADRLMTORCKDKLNALAI SVANQW 120
DB 61 ASGRYEGKITRNSERFELTPNPNPDIIFKDEENTGADRLMTORCKDKLNALAI SVANQW 120
QY 121 PGVKLRYTEGMDGDHSESLHYEGRAVDITTSDDRSKRYGMALRLAVEGPDWYYES 180
DB 121 PGVKLRYTEGMDGDHSESLHYEGRAVDITTSDDRSKRYGMALRLAVEGPDWYYES 180
QY 181 KAHHCYSKAKENSVAAKSGCFPGSATVHLEHGSTKLVKDLSPGDRVLAADADGRLLYSD 240
DB 181 KAHHCYSKAKENSVAAKSGCFPGSATVHLEHGSTKLVKDLSPGDRVLAADADGRLLYSD 240
QY 241 FLTFLDMDSSSKLFLYIETROPARLLLTAAHLLEFVAPOHNOSEATGSGALFASNV 300
DB 241 FLTFLDMDSSSKLFLYIETROPARLLLTAAHLLEFVAPOHNOSEATGSGALFASNV 300
QY 301 KPGORVYVLGEGGOQLLPASVHSVSLREASGAYAPLTAGCTILINVLASCAVIEHS 360
DB 301 KPGORVYVLGEGGOQLLPASVHSVSLREASGAYAPLTAGCTILINVLASCAVIEHS 360
QY 361 MAHMAFAPFLAAGLLAALCPDGAIPTAATTGIIHWSRLYRIGSVLDGDLHPFLGM 420
DB 361 MAHMAFAPFLAAGLLAALCPDGAIPTAATTGIIHWSRLYRIGSVLDGDLHPFLGM 420
QY 421 VAPAS 425
DB 421 VAPAS 425

RESULT 4
US-08-674-509B-8
Sequence 8, Application US/08674509B
Patent No. 6261786
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
APPLICANT: Marijo, Valeria
TITLE OF INVENTION: SCREENING ASSAYS FOR HEDGEHOG AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/674,509B
FILING DATE: 02-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/460,900
FILING DATE: 05-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMV-006.06
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-674-509B-8

Query Match 100.0%; Score 2218; DB 4; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.5e-241;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVEMLLTLLILVGFICALLVSSGLTGCSPGRGIGKRHRPKKLPPLAYKOFIPNVAEKTIG 60
DB 1 MVEMLLTLLILVGFICALLVSSGLTGCSPGRGIGKRHRPKKLPPLAYKOFIPNVAEKTIG 60
QY 61 ASGRYEGKITRNSERFELTPNPNPDIIFKDEENTGADRLMTORCKDKLNALAI SVANQW 120
DB 61 ASGRYEGKITRNSERFELTPNPNPDIIFKDEENTGADRLMTORCKDKLNALAI SVANQW 120
QY 121 PGVKLRYTEGMDGDHSESLHYEGRAVDITTSDDRSKRYGMALRLAVEGPDWYYES 180
DB 121 PGVKLRYTEGMDGDHSESLHYEGRAVDITTSDDRSKRYGMALRLAVEGPDWYYES 180
QY 181 KAHHCYSKAKENSVAAKSGCFPGSATVHLEHGSTKLVKDLSPGDRVLAADADGRLLYSD 240
DB 181 KAHHCYSKAKENSVAAKSGCFPGSATVHLEHGSTKLVKDLSPGDRVLAADADGRLLYSD 240
QY 241 FLTFLDMDSSSKLFLYIETROPARLLLTAAHLLEFVAPOHNOSEATGSGALFASNV 300
DB 241 FLTFLDMDSSSKLFLYIETROPARLLLTAAHLLEFVAPOHNOSEATGSGALFASNV 300
QY 301 KPGORVYVLGEGGOQLLPASVHSVSLREASGAYAPLTAGCTILINVLASCAVIEHS 360
DB 301 KPGORVYVLGEGGOQLLPASVHSVSLREASGAYAPLTAGCTILINVLASCAVIEHS 360
QY 361 MAHMAFAPFLAAGLLAALCPDGAIPTAATTGIIHWSRLYRIGSVLDGDLHPFLGM 420
DB 361 MAHMAFAPFLAAGLLAALCPDGAIPTAATTGIIHWSRLYRIGSVLDGDLHPFLGM 420
QY 421 VAPAS 425
DB 421 VAPAS 425

RESULT 5
US-08-954-698-8
Sequence 8, Application US/08954698
Patent No. 6271363
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing

TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HONG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/954,698
FILING DATE: 20-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/462,386
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/435,093
FILING DATE: 04-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/356,060
FILING DATE: 14-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMV-006.10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-954-698-8

Query Match 100.0%; Score 2218; DB 4; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.5e-241;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWEMLLRILLVGFTICALVSSGLTCGPGRGIGRRRHPKLTPLAYKQFIPNVAEKTIG 60
DB 1 MWEMLLRILLVGFTICALVSSGLTCGPGRGIGRRRHPKLTPLAYKQFIPNVAEKTIG 60
QY 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVMNQ 120
DB 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVMNQ 120
QY 121 PGVKLRVTEGMDEGHSESLHYEGRAVDITTSDRSKYKMLARLAVEAGFDWVYYS 180
DB 121 PGVKLRVTEGMDEGHSESLHYEGRAVDITTSDRSKYKMLARLAVEAGFDWVYYS 180
QY 181 KAHHCQVKAENSVAAKSGGCGPGSATVHLEHGTRKLYKVDLSPGRVLAADADGRLISD 240
DB 181 KAHHCQVKAENSVAAKSGGCGPGSATVHLEHGTRKLYKVDLSPGRVLAADADGRLISD 240
QY 241 FLTFIDRSDSKLFYVETROPARLLTLTAHLTFVAPOHQSEATGSTSGOALFASNV 300
DB 241 FLTFIDRSDSKLFYVETROPARLLTLTAHLTFVAPOHQSEATGSTSGOALFASNV 300
QY 301 KPGQRYVVLGEGGQQLLPASVHVSILREASGAYAPLTAQGITLLINRYLASCYAVIEEHS 360
DB 301 KPGQRYVVLGEGGQQLLPASVHVSILREASGAYAPLTAQGITLLINRYLASCYAVIEEHS 360
QY 361 WAHMAFAPRLAOGILALCPDGAIP7AATTTTG1HMYSRLLYRIGSVWLDGDALHPLGM 420

DB 361 WAHMAFAPRLAOGILALCPDGAIP7AATTTTG1HMYSRLLYRIGSVWLDGDALHPLGM 420
QY 421 VAPAS 425
DB 421 VAPAS 425

RESULT 6
US-08-957-874-8
Sequence 8, Application us/08957874
Patent No. 6384192
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HONG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/957,874
FILING DATE: 20-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/462,386
FILING DATE: 05-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/435,093
FILING DATE: 04-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/356,060
FILING DATE: 14-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMV-006.09
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-957-874-8

Query Match 100.0%; Score 2218; DB 4; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.5e-241;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWEMLLRILLVGFTICALVSSGLTCGPGRGIGRRRHPKLTPLAYKQFIPNVAEKTIG 60
DB 1 MWEMLLRILLVGFTICALVSSGLTCGPGRGIGRRRHPKLTPLAYKQFIPNVAEKTIG 60
QY 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVMNQ 120
DB 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVMNQ 120

QY 121 PGVKLRVTEGMDDEGHSESLHYEGRAVDITTSDRDRSKYGMRLARLAVAGFDWYYES 180
DB 121 PGVKLRVTEGMDDEGHSESLHYEGRAVDITTSDRDRSKYGMRLARLAVAGFDWYYES 180
QY 181 KAHHCVCVKAENSVAAKSGGCFPGSATVHLEHGSTKLYVKLSPEDRYLAADADGRLLYSD 240
DB 181 KAHHCVCVKAENSVAAKSGGCFPGSATVHLEHGSTKLYVKLSPEDRYLAADADGRLLYSD 240
QY 241 FLTFLLDMDSRKLFFYIETROPARLLTLTAHLFLVAPOHNOSEATGTSQALFASNV 300
DB 241 FLTFLLDMDSRKLFFYIETROPARLLTLTAHLFLVAPOHNOSEATGTSQALFASNV 300
QY 301 KPGQRYVVLGEGGQQLLPASVHSVSLREASGAVAPLTAOCTILINRVLASCYAVIEHS 360
DB 301 KPGQRYVVLGEGGQQLLPASVHSVSLREASGAVAPLTAOCTILINRVLASCYAVIEHS 360
QY 361 MAHMAFAPFLAOGLLAALCPDGAIPPAATTTGIIHWYSRLYRIGSWVLDGDLHPLGM 420
DB 361 MAHMAFAPFLAOGLLAALCPDGAIPPAATTTGIIHWYSRLYRIGSWVLDGDLHPLGM 420
QY 421 VAPAS 425
DB 421 VAPAS 425

RESULT 7
US-09-325-256-17
; Sequence 17, Application US/09325256
; Patent No. 6444793
; GENERAL INFORMATION:
; APPLICANT: PEPINSKY, R. BLAKE
; APPLICANT: BAKER, DARREN P.
; APPLICANT: MEN, DINGYI
; APPLICANT: WILLIAMS, KEVIN P.
; APPLICANT: GARGER, ELLEN A.
; APPLICANT: TAYLOR, FREDERICK R.
; APPLICANT: GALDES, ALPHONSE
; APPLICANT: PORTER, JEFFREY
; TITLE OF INVENTION: HYDROPHOBICALLY-MODIFIED PROTEIN COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; FILE REFERENCE: BIV-067-01
; CURRENT APPLICATION NUMBER: US/09/325,256
; CURRENT FILING DATE: 1999-06-03
; PRIOR APPLICATION NUMBER: 60/099,800
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/078,935
; PRIOR FILING DATE: 1998-03-20
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089,685
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/067,423
; PRIOR FILING DATE: 1997-12-03
; PRIOR APPLICATION NUMBER: PCT/US98/25676
; PRIOR FILING DATE: 1998-12-03
; NUMBER OF SEQ. ID NOS: 31
; SOFTWARE: Patentln Ver. 2.1
; SEQ. ID NO: 17
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Gallus sp.
US-09-325-256-17

Query Match 100.0%; Score 2218; DB 4; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.5e-241;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVEKLLTRILLVGFALLVSSSLTGPRGIGTKRRHPKLTPLAKOKTIPNVAETKLG 60
DB 1 MVEKLLTRILLVGFALLVSSSLTGPRGIGTKRRHPKLTPLAKOKTIPNVAETKLG 60
QY 61 ASGRYBCKITRNSERFKELTPNYPDITTFKDEENTGADRLMTORCKKLAALATSVNOM 120
DB 61 ASGRYBCKITRNSERFKELTPNYPDITTFKDEENTGADRLMTORCKKLAALATSVNOM 120

QY 121 PGVKLRVTEGMDDEGHSESLHYEGRAVDITTSDRDRSKYGMRLARLAVAGFDWYYES 180
DB 121 PGVKLRVTEGMDDEGHSESLHYEGRAVDITTSDRDRSKYGMRLARLAVAGFDWYYES 180
QY 181 KAHHCVCVKAENSVAAKSGGCFPGSATVHLEHGSTKLYVKLSPEDRYLAADADGRLLYSD 240
DB 181 KAHHCVCVKAENSVAAKSGGCFPGSATVHLEHGSTKLYVKLSPEDRYLAADADGRLLYSD 240
QY 241 FLTFLLDMDSRKLFFYIETROPARLLTLTAHLFLVAPOHNOSEATGTSQALFASNV 300
DB 241 FLTFLLDMDSRKLFFYIETROPARLLTLTAHLFLVAPOHNOSEATGTSQALFASNV 300
QY 301 KPGQRYVVLGEGGQQLLPASVHSVSLREASGAVAPLTAOCTILINRVLASCYAVIEHS 360
DB 301 KPGQRYVVLGEGGQQLLPASVHSVSLREASGAVAPLTAOCTILINRVLASCYAVIEHS 360
QY 361 MAHMAFAPFLAOGLLAALCPDGAIPPAATTTGIIHWYSRLYRIGSWVLDGDLHPLGM 420
DB 361 MAHMAFAPFLAOGLLAALCPDGAIPPAATTTGIIHWYSRLYRIGSWVLDGDLHPLGM 420
QY 421 VAPAS 425
DB 421 VAPAS 425

RESULT 8
US-08-946-329A-19
; Sequence 19, Application US/08946329A
; Patent No. 6057091
; GENERAL INFORMATION:
; APPLICANT: Beachy, Philip A.
; APPLICANT: Porter, Jeffrey A.
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
; NUMBER OF SEQUENCES: 109
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/946,329A
; FILING DATE: 07-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/061,323
; FILING DATE: 07-OCT-1996
; APPLICATION NUMBER: 08/729,743
; FILING DATE: 10-JUL-1996
; APPLICATION NUMBER: 08/567,357
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/349,498
; FILING DATE: 02-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/140001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ. ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: protein
US-08-946-329A-19

Query Match 94.0%; Score 2085; DB 3; Length 425;
Best Local Similarity 96.2%; Pred. No. 1.6e-226;
Matches 410; Conservative 1; Mismatches 13; Indels 2; Gaps 2;

QY 1 MEMLLRLLVGFICALVSSGLTCGPGRGIGRRHPPKLTPLAYKQFIPIVNAEKTG 60
DB 1 MEMLLRLLVGFICALVSSGLTCGPGRGIGRRHPPKLTPLAYKQFIPIVNAEKTG 60
QY 61 ASGREGKTRNSERFKELIPYNPDIIFKDEENTGADRLMTORCKDKLNALATISVMNQ 120
DB 61 ASGREGKTRNSERFKELIPYNPDIIFKDEENTGADRLMTORCKDKLNALATISVMNQ 120
QY 121 PGVKLRVTEGMEDEGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDVYYES 180
DB 121 PGVKLRVTEGMEDEGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDVYYES 180
QY 181 KAHICSVKAENSVAKSGCGPFSATYHLEHGTRKLYKDLSPGDRVLAADADGRLLVSD 240
DB 181 KAHICSVKAENSVAKSGCGPFSATYHLEHGTRKLYKDLSPGDRVLAADADGRLLVSD 240
QY 241 F-LTFLDRMDSRRKLFYIETROPARALLTAHLLFVAPQHOSEATGTSGLAFASN 299
DB 240 FLTLFLDRMDSRRKLFYIETROPARALLTAHLLFVAPQHOSEATGTSGLAFASN 299
QY 300 VPGQRYVYLGEQGQQLLPASVHSVLSREASGAYAPLTAOGTILINRVLASCVAVIEEH 359
DB 300 VPGQRYVYLGEQGQQLLPASVHSVLSREASGAYAPLTAOGTILINRVLASCVAVIEEH 359
QY 360 SWAHNAFAPFRLAOGLLALCPDGAIPTAATTTGIIHWSRLYRIGSVWLDGDLHPLG 419
DB 360 SWAHNAFAPFRLAOGLLALCPDGAIPTAATTTGIIHWSRLYRIGSVWLDGDLHPLG 419
QY 420 WYAPAS 425
DB 420 WYAPAS 425

RESULT 9
US-08-567-357A-19
; Sequence 19, Application US/08567357A
; Patent No. 6132728
; GENERAL INFORMATION:
; APPLICANT: Beachy, Philip A.
; APPLICANT: Moon, Randall T.
; APPLICANT: Porter, Jeffrey A.
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/567,357A
; FILING DATE: 04-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/349,498
; FILING DATE: 02-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/080001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: protein
US-08-567-357A-19

Query Match 94.0%; Score 2085; DB 4; Length 425;
Best Local Similarity 96.2%; Pred. No. 1.6e-226;
Matches 410; Conservative 1; Mismatches 13; Indels 2; Gaps 2;

QY 1 MEMLLRLLVGFICALVSSGLTCGPGRGIGRRHPPKLTPLAYKQFIPIVNAEKTG 60
DB 1 MEMLLRLLVGFICALVSSGLTCGPGRGIGRRHPPKLTPLAYKQFIPIVNAEKTG 60
QY 61 ASGREGKTRNSERFKELIPYNPDIIFKDEENTGADRLMTORCKDKLNALATISVMNQ 120
DB 61 ASGREGKTRNSERFKELIPYNPDIIFKDEENTGADRLMTORCKDKLNALATISVMNQ 120
QY 121 PGVKLRVTEGMEDEGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDVYYES 180
DB 121 PGVKLRVTEGMEDEGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDVYYES 180
QY 181 KAHICSVKAENSVAKSGCGPFSATYHLEHGTRKLYKDLSPGDRVLAADADGRLLVSD 240
DB 181 KAHICSVKAENSVAKSGCGPFSATYHLEHGTRKLYKDLSPGDRVLAADADGRLLVSD 240
QY 241 F-LTFLDRMDSRRKLFYIETROPARALLTAHLLFVAPQHOSEATGTSGLAFASN 299
DB 240 FLTLFLDRMDSRRKLFYIETROPARALLTAHLLFVAPQHOSEATGTSGLAFASN 299
QY 300 VPGQRYVYLGEQGQQLLPASVHSVLSREASGAYAPLTAOGTILINRVLASCVAVIEEH 359
DB 300 VPGQRYVYLGEQGQQLLPASVHSVLSREASGAYAPLTAOGTILINRVLASCVAVIEEH 359
QY 360 SWAHNAFAPFRLAOGLLALCPDGAIPTAATTTGIIHWSRLYRIGSVWLDGDLHPLG 419
DB 360 SWAHNAFAPFRLAOGLLALCPDGAIPTAATTTGIIHWSRLYRIGSVWLDGDLHPLG 419
QY 420 WYAPAS 425
DB 420 WYAPAS 425

RESULT 10
US-08-729-743A-19
; Sequence 19, Application US/08729743A
; Patent No. 6214794
; GENERAL INFORMATION:
; APPLICANT: Beachy, Philip A.
; APPLICANT: Moon, Randall T.
; APPLICANT: Porter, Jeffrey A.
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,743A
; FILING DATE: 07-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/567,357
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/349,498

FILING DATE: 02-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Hallie, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/099001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
TOPOLOGY: both
MOLECULE TYPE: protein
US-08-729-743A-19

Query Match 94.0%; Score 2085; DB 4; Length 425;
Best Local Similarity 96.2%; Pred. No.1.6e-226;
Matches 410; Conservative 1; Mismatches 13; Indels 2; Gaps 2;

QY 1 MVELLTLRLVGFICALVSSGLTCGPGRGIGKRRHPKRLPLAYKQFIPNVAEKTG 60
DB 1 MVELLTLRLVGFICALVSSGLTCGPGRGIGKRRHPKRLPLAYKQFIPNVAEKTG 60
QY 61 ASGRYEGKITRNSERKELTPNTNPDIIFKDEENTGADRLMTORCKDKLNALISVNCW 120
DB 61 ASGRYEGKITRNSERKELTPNTNPDIIFKDEENTGADRLMTORCKDKLNALISVNCW 120
QY 121 PGVLRTEGMDGHHSESLHYEGRAVDITSDRDSKYYGMARLAVAGDWMYVES 180
DB 121 PGVLRTEGMDGHHSESLHYEGRAVDITSDRDSKYYGMARLAVAGDWMYVES 180
QY 181 KAHHSVKAENSVAAKSGCFFGSATVHLEHGKTKLVKDLSPGDRVLADADGRLLYSD 240
DB 181 KAHHSVKAENSVAAKSGCFFGSATVHLEHGKTKLVKDLSPGDRVLADADGRLLYSD 240
QY 241 F-LTFDRLDMSRKLFYVETROPARLLTAAHLFVAPQHNOSEATGSGQALFASN 299
DB 241 F-LTFDRLDMSRKLFYVETROPARLLTAAHLFVAPQHNOSEATGSGQALFASN 299
QY 300 VKPGORVYVLEGGOQLLPASVHSVSLREBSAGVAPLTAGTLLINRYLASCYAVTEEH 359
DB 300 VKPGORVYVLEGGOQLLPASVHSVSLREBSAGVAPLTAGTLLINRYLASCYAVTEEH 359
QY 360 SMAHMAFAPRLAOGILAAICPDGAIPTAATTTGIMYSRLLYRIGSWVLDGDLHPLG 419
DB 360 SMAHMAFAPRLAOGILAAICPDGAIPTAATTTGIMYSRLLYRIGSWVLDGDLHPLG 419
QY 420 MVAAPAS 425
DB 420 MVAAPAS 425

RESULT 11
US-08-349-498-19
Sequence 19, Application US/08349498
Patent No. 6281332
GENERAL INFORMATION:
APPLICANT: Beachy, Philip A.
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: U.S.A.
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/349,498
FILING DATE: 02-DEC-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Hallie, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/043001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: both
MOLECULE TYPE: protein
US-08-349-498-19

Query Match 94.0%; Score 2085; DB 4; Length 425;
Best Local Similarity 96.2%; Pred. No.1.6e-226;
Matches 410; Conservative 1; Mismatches 13; Indels 2; Gaps 2;

QY 1 MVELLTLRLVGFICALVSSGLTCGPGRGIGKRRHPKRLPLAYKQFIPNVAEKTG 60
DB 1 MVELLTLRLVGFICALVSSGLTCGPGRGIGKRRHPKRLPLAYKQFIPNVAEKTG 60
QY 61 ASGRYEGKITRNSERKELTPNTNPDIIFKDEENTGADRLMTORCKDKLNALISVNCW 120
DB 61 ASGRYEGKITRNSERKELTPNTNPDIIFKDEENTGADRLMTORCKDKLNALISVNCW 120
QY 121 PGVLRTEGMDGHHSESLHYEGRAVDITSDRDSKYYGMARLAVAGDWMYVES 180
DB 121 PGVLRTEGMDGHHSESLHYEGRAVDITSDRDSKYYGMARLAVAGDWMYVES 180
QY 181 KAHHSVKAENSVAAKSGCFFGSATVHLEHGKTKLVKDLSPGDRVLADADGRLLYSD 240
DB 181 KAHHSVKAENSVAAKSGCFFGSATVHLEHGKTKLVKDLSPGDRVLADADGRLLYSD 240
QY 241 F-LTFDRLDMSRKLFYVETROPARLLTAAHLFVAPQHNOSEATGSGQALFASN 299
DB 241 F-LTFDRLDMSRKLFYVETROPARLLTAAHLFVAPQHNOSEATGSGQALFASN 299
QY 300 VKPGORVYVLEGGOQLLPASVHSVSLREBSAGVAPLTAGTLLINRYLASCYAVTEEH 359
DB 300 VKPGORVYVLEGGOQLLPASVHSVSLREBSAGVAPLTAGTLLINRYLASCYAVTEEH 359
QY 360 SMAHMAFAPRLAOGILAAICPDGAIPTAATTTGIMYSRLLYRIGSWVLDGDLHPLG 419
DB 360 SMAHMAFAPRLAOGILAAICPDGAIPTAATTTGIMYSRLLYRIGSWVLDGDLHPLG 419
QY 420 MVAAPAS 425
DB 420 MVAAPAS 425

RESULT 12
PCT-US95-15463-19
Sequence 19, Application PC/TUS9515463
GENERAL INFORMATION:
APPLICANT: The Johns Hopkins University School of Medicine
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: U.S.A.
ZIP: 92037
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15463
FILING DATE: 01-DEC-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/080M01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: both
MOLECULE TYPE: protein
PCT-US95-15463-19

Query Match 94.0%; Score 2085; DB 5; Length 425;
Best Local Similarity 96.2%; Pred. No. 1.6e-226;
Matches 410; Conservative 1; Mismatches 13; Indels 2; Gaps 2;

QY 1 MVEMLLTLLVGFICALVSSGLTCGPGRGIGRRHHPKLTPLAYKOFINVAEKTIG 60
DB 1 MVEMLLTLLVGFICALVSSGLTCGPGRGIGRRHHPKLTPLAYKOFINVAEKTIG 60
QY 61 ASGREGKTRNSERKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAIIVMOW 120
DB 61 ASGREGKTRNSERKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAIIVMOW 120
QY 121 PGVKLRVTGMDGDHSHSESLHYEGRAVDITTSDBDSKTYMLRLAVEAGFDVYYES 180
DB 121 PGVKLRVTGMDGDHSHSESLHYEGRAVDITTSDBDSKTYMLRLAVEAGFDVYYES 180
QY 181 KAHICSVKAENSVAAKSGCGPFSATVHLEHGRTKLVKDSLPGDRLVLAADGRLVSD 240
DB 181 KAHICSVKAENSVAAKSGCGPFSATVHLEHGRTKLVKDSLPGDRLVLAADGRLVSD 240
QY 241 F-LTFLDRDSSRKLEFYIETROPARALLTAAHLLFVAPQHNOSEATGSTSGALFASN 299
DB 241 F-LTFLDRDSSRKLEFYIETROPARALLTAAHLLFVAPQHNOSEATGSTSGALFASN 299
QY 300 VPRGQRYVVLGGGQOLLPASVHSVSLREASGATAPLTAOCTILINRYLASCYAVIEEH 359
DB 300 VPRGQRYVVLGGGQOLLPASVHSVSLREASGATAPLTAOCTILINRYLASCYAVIEEH 359
QY 360 SWAHNAFAPFRLAOGILALCPDGAIPPTAATTTGIIHWSRLLYRIGSVLDGDLHPUG 419
DB 360 SWAHNAFAPFRLAOGILALCPDGAIPPTAATTTGIIHWSRLLYRIGSVLDGDLHPUG 419
QY 420 MYAPAS 425
DB 420 MYAPAS 425

RESULT 13
PCT-US95-15923-19
Sequence 19, Application PC/TUS9515923
GENERAL INFORMATION:
APPLICANT: The Johns Hopkins University School of Medicine, et al.
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA

COUNTRY: U.S.A.
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15923
FILING DATE: 04-DEC-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/043M01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: both
MOLECULE TYPE: protein
PCT-US95-15923-19

Query Match 94.0%; Score 2085; DB 5; Length 425;
Best Local Similarity 96.2%; Pred. No. 1.6e-226;
Matches 410; Conservative 1; Mismatches 13; Indels 2; Gaps 2;

QY 1 MVEMLLTLLVGFICALVSSGLTCGPGRGIGRRHHPKLTPLAYKOFINVAEKTIG 60
DB 1 MVEMLLTLLVGFICALVSSGLTCGPGRGIGRRHHPKLTPLAYKOFINVAEKTIG 60
QY 61 ASGREGKTRNSERKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAIIVMOW 120
DB 61 ASGREGKTRNSERKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAIIVMOW 120
QY 121 PGVKLRVTGMDGDHSHSESLHYEGRAVDITTSDBDSKTYMLRLAVEAGFDVYYES 180
DB 121 PGVKLRVTGMDGDHSHSESLHYEGRAVDITTSDBDSKTYMLRLAVEAGFDVYYES 180
QY 181 KAHICSVKAENSVAAKSGCGPFSATVHLEHGRTKLVKDSLPGDRLVLAADGRLVSD 240
DB 181 KAHICSVKAENSVAAKSGCGPFSATVHLEHGRTKLVKDSLPGDRLVLAADGRLVSD 240
QY 241 F-LTFLDRDSSRKLEFYIETROPARALLTAAHLLFVAPQHNOSEATGSTSGALFASN 299
DB 241 F-LTFLDRDSSRKLEFYIETROPARALLTAAHLLFVAPQHNOSEATGSTSGALFASN 299
QY 300 VPRGQRYVVLGGGQOLLPASVHSVSLREASGATAPLTAOCTILINRYLASCYAVIEEH 359
DB 300 VPRGQRYVVLGGGQOLLPASVHSVSLREASGATAPLTAOCTILINRYLASCYAVIEEH 359
QY 360 SWAHNAFAPFRLAOGILALCPDGAIPPTAATTTGIIHWSRLLYRIGSVLDGDLHPUG 419
DB 360 SWAHNAFAPFRLAOGILALCPDGAIPPTAATTTGIIHWSRLLYRIGSVLDGDLHPUG 419
QY 420 MYAPAS 425
DB 420 MYAPAS 425

RESULT 14
US-08-748-591-4
Sequence 4, Application US/08748591
Patent No. 5759811
GENERAL INFORMATION:
APPLICANT: Epstein, Ervin
APPLICANT: Hu, Zhilan
APPLICANT: Bonifas, Jeanette
TITLE OF INVENTION: Mutant Human Hedgehog Gene

NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish and Richardson
STREET: 2200 Sand Hill Road
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Releasee #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/748,591
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: 06510/067001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 322-5070
TELEFAX: (415) 854-0875
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-748-591-4

Query Match 81.5%; Score 1807; DB 1; Length 462;
Best Local Similarity 77.8%; Pred. No. 4.7e-195;
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;
QY 4 MLLTRILLVGFICALVSSGLTCGPGRGIGKRRHPKLTPLAVYKQPIPVNAEKTIGASG 63
DB 1 MLLARCLLVIVSSLLVCSGLACGPGRGIGKRRHPKLTPLAVYKQPIPVNAEKTIGASG 60
QY 64 RYEGKITRNSERKELTPNNNPDIIFKDEENTGADRLMTORCKDKLNALISVNMOPGV 123
DB 61 RYEGKITRNSERKELTPNNNPDIIFKDEENTGADRLMTORCKDKLNALISVNMOPGV 120
QY 124 KLVTEGMDSDGHSESLAYEGRAVDITTSDDRKYGMILARLAVAGFDWYYESKAH 183
DB 121 KLVTEGMDSDGHSESLAYEGRAVDITTSDDRKYGMILARLAVAGFDWYYESKAH 180
QY 184 IHCYKAENSVAASKGCGPFGSATVHLEHGTRKLVKDLSPGDRVLAADADGRLLYSDFLT 243
DB 181 IHCYKAENSVAASKGCGPFGSATVHLEHGTRKLVKDLSPGDRVLAADADGRLLYSDFLT 240
QY 244 FLDRMDSRKLFFVYIETROPARLLTLAHLLEVAPOHNOSEATG-----STSG----- 292
DB 241 FLDRMDSRKLFFVYIETROPARLLTLAHLLEVAPOHNOSEATG-----STSG----- 292
QY 293 ---QALFASNVKPGORYVGE---GGQQLPASVHSVSLREASGAVAPLTAGCTIILNR 347
DB 299 LGPRALFASNVKPGORYVGE---GGQQLPASVHSVSLREASGAVAPLTAGCTIILNR 347
QY 348 VLASCVAVIEHSHMAHAFPFRLAHLAALPARTDRGDSGGGGRGGRVALTAP 418
DB 359 VLASCVAVIEHSHMAHAFPFRLAHLAALPARTDRGDSGGGGRGGRVALTAP 418
QY 382 DGAIPATAATTTGIIHWYSRLLYRIGSVNLDGDLHPGLGAVAPAS 425
DB 419 GADAPGAGATAGIHWYSRLLYIGTWLDSALHPLGMAVKSS 462

RESULT 15
US-08-748-591-9
; Sequence 9, Application US/08748591

Patent No. 5759811
GENERAL INFORMATION:
APPLICANT: Epstein, Ervin
APPLICANT: Hu, Zhilan
APPLICANT: Bonifas, Jeanette
TITLE OF INVENTION: Mutant Human Hedgehog Gene
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish and Richardson
STREET: 2200 Sand Hill Road
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Releasee #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/748,591
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: 06510/067001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 322-5070
TELEFAX: (415) 854-0875
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-748-591-9

Query Match 81.5%; Score 1807; DB 1; Length 462;
Best Local Similarity 77.8%; Pred. No. 4.7e-195;
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;
QY 4 MLLTRILLVGFICALVSSGLTCGPGRGIGKRRHPKLTPLAVYKQPIPVNAEKTIGASG 63
DB 1 MLLARCLLVIVSSLLVCSGLACGPGRGIGKRRHPKLTPLAVYKQPIPVNAEKTIGASG 60
QY 64 RYEGKITRNSERKELTPNNNPDIIFKDEENTGADRLMTORCKDKLNALISVNMOPGV 123
DB 61 RYEGKITRNSERKELTPNNNPDIIFKDEENTGADRLMTORCKDKLNALISVNMOPGV 120
QY 124 KLVTEGMDSDGHSESLAYEGRAVDITTSDDRKYGMILARLAVAGFDWYYESKAH 183
DB 121 KLVTEGMDSDGHSESLAYEGRAVDITTSDDRKYGMILARLAVAGFDWYYESKAH 180
QY 184 IHCYKAENSVAASKGCGPFGSATVHLEHGTRKLVKDLSPGDRVLAADADGRLLYSDFLT 243
DB 181 IHCYKAENSVAASKGCGPFGSATVHLEHGTRKLVKDLSPGDRVLAADADGRLLYSDFLT 240
QY 244 FLDRMDSRKLFFVYIETROPARLLTLAHLLEVAPOHNOSEATG-----STSG----- 292
DB 241 FLDRMDSRKLFFVYIETROPARLLTLAHLLEVAPOHNOSEATG-----STSG----- 292
QY 293 ---QALFASNVKPGORYVGE---GGQQLPASVHSVSLREASGAVAPLTAGCTIILNR 347
DB 299 LGPRALFASNVKPGORYVGE---GGQQLPASVHSVSLREASGAVAPLTAGCTIILNR 347
QY 348 VLASCVAVIEHSHMAHAFPFRLAHLAALPARTDRGDSGGGGRGGRVALTAP 418
DB 359 VLASCVAVIEHSHMAHAFPFRLAHLAALPARTDRGDSGGGGRGGRVALTAP 418
QY 382 DGAIPATAATTTGIIHWYSRLLYRIGSVNLDGDLHPGLGAVAPAS 425

Thu Feb 20 10:43:48 2003

us-09-827-110a-10.ra1

Page 10

DB 419 GAADAFGAGATAGIHWYSOLLYQIGTWLIDSEALHPLGMAYKSS 462

Search completed: February 20, 2003, 10:14:10
Job time : 13.1257 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Comphen Ltd.

OM protein - protein search, using SW model

Run on: February 20, 2003, 10:10:28 ; Search time 7.62902 Seconds
(without alignments)
1423.285 Million cell updates/sec

Title: US-09-827-110a-10

Sequence: 1 MVEMLLRILLVGFICAL.....GSWLDGDLHPLGVAPAS 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 2554876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2218	100.0	425	US-08-900-220C-10	Sequence 10, Appl
2	2218	100.0	425	US-09-883-848A-10	Sequence 10, Appl
3	2218	100.0	425	US-09-021-660A-34	Sequence 34, Appl
4	2218	100.0	425	US-09-151-999-10	Sequence 10, Appl
5	1807	81.5	462	US-09-733-634-14	Sequence 14, Appl
6	1807	81.5	475	US-08-900-220C-15	Sequence 15, Appl
7	1807	81.5	475	US-09-883-848A-15	Sequence 15, Appl
8	1807	81.5	475	US-09-021-660A-39	Sequence 39, Appl
9	1807	81.5	475	US-09-151-999-15	Sequence 15, Appl
10	1791	80.7	437	US-09-990-046-14	Sequence 14, Appl
11	1791	80.7	437	US-09-969-520A-1	Sequence 1, Appl
12	1791	80.7	437	US-10-013-310-1	Sequence 1, Appl
13	1791	80.7	437	US-09-733-634-16	Sequence 16, Appl
14	1791	80.7	437	US-09-021-660A-37	Sequence 37, Appl
15	1788	80.6	437	US-08-900-220C-13	Sequence 13, Appl
16	1788	80.6	437	US-09-883-848A-13	Sequence 13, Appl
17	1788	80.6	437	US-09-151-999-13	Sequence 13, Appl
18	1785	80.5	437	US-09-969-520A-3	Sequence 3, Appl
19	1782	80.3	437	US-09-969-520A-10	Sequence 10, Appl

20	1782	80.3	437	US-09-969-520A-11	Sequence 11, Appl
21	1781	80.3	437	US-09-969-520A-2	Sequence 2, Appl
22	1777	80.1	437	US-09-969-520A-9	Sequence 9, Appl
23	1775	80.0	437	US-09-969-520A-4	Sequence 4, Appl
24	1771	79.8	437	US-09-969-520A-8	Sequence 8, Appl
25	1768	79.7	437	US-09-969-520A-6	Sequence 6, Appl
26	1759	79.3	437	US-09-969-520A-7	Sequence 7, Appl
27	1742	78.5	437	US-09-969-520A-5	Sequence 5, Appl
28	1494	67.4	418	US-09-021-660A-38	Sequence 38, Appl
29	1458	65.7	418	US-08-900-220C-18	Sequence 18, Appl
30	1458	65.7	416	US-09-883-848A-18	Sequence 18, Appl
31	1458	65.7	416	US-09-151-999-18	Sequence 18, Appl
32	1430	64.5	418	US-08-900-220C-14	Sequence 14, Appl
33	1430	64.5	418	US-09-883-848A-14	Sequence 14, Appl
34	1430	64.5	418	US-09-151-999-14	Sequence 14, Appl
35	1277.5	57.6	449	US-09-990-046-29	Sequence 29, Appl
36	1277.5	57.6	449	US-09-733-634-30	Sequence 30, Appl
37	1270.5	57.3	411	US-08-900-220C-16	Sequence 16, Appl
38	1270.5	57.3	411	US-09-883-848A-16	Sequence 16, Appl
39	1270.5	57.3	411	US-09-151-999-16	Sequence 16, Appl
40	1265.5	57.1	411	US-09-733-634-28	Sequence 28, Appl
41	1249.5	56.3	411	US-08-900-220C-12	Sequence 12, Appl
42	1249.5	56.3	411	US-09-883-848A-12	Sequence 12, Appl
43	1249.5	56.3	411	US-09-151-999-12	Sequence 12, Appl
44	1200	54.1	396	US-08-900-220C-11	Sequence 11, Appl
45	1200	54.1	396	US-09-990-046-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-900-220C-10
Sequence 10, Application US/08900220C
Patent No. US20020045206A1
GENERAL INFORMATION:
APPLICANT: Miao, Ningning
Wang, Manhappa
Pang, Kevin
Jin, Ping
TITLE OF INVENTION: Method of Treating Dopaminergic and GABA-nergic Disorders
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSER: FOLEY, HOAG & ELLIOT LLP
STREET: ONE POST OFFICE SQUARE
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/900,220C
FILING DATE: 24-Jul-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: ONV-044.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-08-900-220C-10

Query Match 100.0%; Score 2218; DB 8; Length 425;
Best Local Similarity 100.0%; Pred. No. 7.8e-196;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WEMLLLRLLLVGFCALVSSGLTCGPGRGIGRRRHPKLTPLAYKQFIPNVAEKTIG 60
DB 1 WEMLLLRLLLVGFCALVSSGLTCGPGRGIGRRRHPKLTPLAYKQFIPNVAEKTIG 60
QY 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDLNALAISVMNQ 120
DB 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDLNALAISVMNQ 120
QY 121 PGVKLRVTEGMEDEGHSEESLHYEGRAVDITTSDRDSKYGMLARLAVEAGFDWYTES 180
DB 121 PGVKLRVTEGMEDEGHSEESLHYEGRAVDITTSDRDSKYGMLARLAVEAGFDWYTES 180
QY 181 KAHICSVKAENSVAAKSGCGPESATVHLEHGTRKLVKDLSPGDRVLAADADGRILYSD 240
DB 181 KAHICSVKAENSVAAKSGCGPESATVHLEHGTRKLVKDLSPGDRVLAADADGRILYSD 240
QY 241 FLTFIDRMDSSRKLFYVETROPARLLTLTAHLFVAPOHNSATGSTSGQALFASNV 300
DB 241 FLTFIDRMDSSRKLFYVETROPARLLTLTAHLFVAPOHNSATGSTSGQALFASNV 300
QY 301 KPGQRYVYLGEQGQQLLPASVHSVSLREBASGAVAPLTAOGTILINRVLASCYAVIEHS 360
DB 301 KPGQRYVYLGEQGQQLLPASVHSVSLREBASGAVAPLTAOGTILINRVLASCYAVIEHS 360
QY 361 WAHMAFAPFRLAOGLLALCPDGAIPTAATTTTGTHWYSRLLYRIGSVLGDALHPHGM 420
DB 361 WAHMAFAPFRLAOGLLALCPDGAIPTAATTTTGTHWYSRLLYRIGSVLGDALHPHGM 420
QY 421 VAPAS 425
DB 421 VAPAS 425

RESULT 2
US-09-883-848A-10
Sequence 10, Application US/09883848A
Publication No. US20030022819A1
GENERAL INFORMATION:
APPLICANT: Ling, L.
APPLICANT: Sanicola-Nadel, M.
TITLE OF INVENTION: ANGIOGENESIS-MODULATING COMPOSITIONS AND USES
FILE REFERENCE: CIBT-P01-119
CURRENT APPLICATION NUMBER: US/09/883, 848A
CURRENT FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/211, 919
PRIOR FILING DATE: 2000-06-16
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 425
TYPE: PRT
ORGANISM: Gallus gallus
US-09-883-848A-10

Query Match 100.0%; Score 2218; DB 9; Length 425;
Best Local Similarity 100.0%; Pred. No. 7.8e-196;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WEMLLLRLLLVGFCALVSSGLTCGPGRGIGRRRHPKLTPLAYKQFIPNVAEKTIG 60
DB 1 WEMLLLRLLLVGFCALVSSGLTCGPGRGIGRRRHPKLTPLAYKQFIPNVAEKTIG 60
QY 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDLNALAISVMNQ 120
DB 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDLNALAISVMNQ 120

QY 121 PGVKLRVTEGMEDEGHSEESLHYEGRAVDITTSDRDSKYGMLARLAVEAGFDWYTES 180
DB 121 PGVKLRVTEGMEDEGHSEESLHYEGRAVDITTSDRDSKYGMLARLAVEAGFDWYTES 180
QY 181 KAHICSVKAENSVAAKSGCGPESATVHLEHGTRKLVKDLSPGDRVLAADADGRILYSD 240
DB 181 KAHICSVKAENSVAAKSGCGPESATVHLEHGTRKLVKDLSPGDRVLAADADGRILYSD 240
QY 241 FLTFIDRMDSSRKLFYVETROPARLLTLTAHLFVAPOHNSATGSTSGQALFASNV 300
DB 241 FLTFIDRMDSSRKLFYVETROPARLLTLTAHLFVAPOHNSATGSTSGQALFASNV 300
QY 301 KPGQRYVYLGEQGQQLLPASVHSVSLREBASGAVAPLTAOGTILINRVLASCYAVIEHS 360
DB 301 KPGQRYVYLGEQGQQLLPASVHSVSLREBASGAVAPLTAOGTILINRVLASCYAVIEHS 360
QY 361 WAHMAFAPFRLAOGLLALCPDGAIPTAATTTTGTHWYSRLLYRIGSVLGDALHPHGM 420
DB 361 WAHMAFAPFRLAOGLLALCPDGAIPTAATTTTGTHWYSRLLYRIGSVLGDALHPHGM 420
QY 421 VAPAS 425
DB 421 VAPAS 425

RESULT 3
US-09-021-660A-34
Sequence 34, Application US/09021660A
Patent No. US20010041668A1
GENERAL INFORMATION:
APPLICANT: Baron, M.
APPLICANT: Belaussoff, M.
TITLE OF INVENTION: METHODS FOR MODULATING HEMATOPOIESIS AND VASCULAR
TITLE OF INVENTION: GROWTH
FILE REFERENCE: H01P-P01-060
CURRENT APPLICATION NUMBER: US/09/021, 660A
CURRENT FILING DATE: 2001-08-27
PRIOR APPLICATION NUMBER: 60/037, 513
PRIOR FILING DATE: 1997-02-10
PRIOR APPLICATION NUMBER: 60/049, 763
PRIOR FILING DATE: 1997-06-16
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 34
LENGTH: 425
TYPE: PRT
ORGANISM: Gallus gallus
US-09-021-660A-34

Query Match 100.0%; Score 2218; DB 10; Length 425;
Best Local Similarity 100.0%; Pred. No. 7.8e-196;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WEMLLLRLLLVGFCALVSSGLTCGPGRGIGRRRHPKLTPLAYKQFIPNVAEKTIG 60
DB 1 WEMLLLRLLLVGFCALVSSGLTCGPGRGIGRRRHPKLTPLAYKQFIPNVAEKTIG 60
QY 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDLNALAISVMNQ 120
DB 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDLNALAISVMNQ 120
QY 121 PGVKLRVTEGMEDEGHSEESLHYEGRAVDITTSDRDSKYGMLARLAVEAGFDWYTES 180
DB 121 PGVKLRVTEGMEDEGHSEESLHYEGRAVDITTSDRDSKYGMLARLAVEAGFDWYTES 180
QY 181 KAHICSVKAENSVAAKSGCGPESATVHLEHGTRKLVKDLSPGDRVLAADADGRILYSD 240
DB 181 KAHICSVKAENSVAAKSGCGPESATVHLEHGTRKLVKDLSPGDRVLAADADGRILYSD 240
QY 241 FLTFIDRMDSSRKLFYVETROPARLLTLTAHLFVAPOHNSATGSTSGQALFASNV 300
DB 241 FLTFIDRMDSSRKLFYVETROPARLLTLTAHLFVAPOHNSATGSTSGQALFASNV 300


```

1 STATE: MA
2 COUNTRY: USA
3 ZIP: 02109
4
5 COMPUTER READABLE FORM:
6 MEDIUM TYPE: Floppy disk
7 COMPUTER: IBM PC compatible
8 OPERATING SYSTEM: PC-DOS/MS-DOS
9 SOFTWARE: Asclit (text)
10
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/08/900,220C
13 FILING DATE: 24-Jul-1997
14 CLASSIFICATION: <unknown>
15 ATTORNEY/AGENT INFORMATION:
16 NAME: Vincent, Matthew P.
17 REGISTRATION NUMBER: 36,709
18 REFERENCE/DOCKET NUMBER: ONV-044.01
19 TELECOMMUNICATION INFORMATION:
20 TELEPHONE: (617) 832-1000
21 TELEFAX: (617) 832-7000
22
23 INFORMATION FOR SEQ ID NO: 15:
24 SEQUENCE CHARACTERISTICS:
25 LENGTH: 475 amino acids
26 TYPE: amino acid
27 TOPOLOGY: linear
28 MOLECULE TYPE: protein
29 SEQUENCE DESCRIPTION: SEQ ID NO: 15:
30
31 US-08-900-220C-15

```

```

; SEQ ID NO 39
; LENGTH: 475
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (463)
; OTHER INFORMATION: Xaa-unknown amino acid
US-09-021-660A-39

```

```

Query Match      81.5%; Score 1807; DB 10; Length 475;
Best Local Similarity 77.8%; Pred. No. 5.3e-158;
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;

```

```

QY 4 MLLTTRLLVGFICALVSSGLTCGPGRGIGKRRHPKRLPLAYKQFIPVNAEKTIGASC 63
   1 MLLARCLLVLVSSLLVCSGLACGPGRGIGKRRHPKRLPLAYKQFIPVNAEKTIGASC 60
DB 1
QY 64 RYEGKITRNSERKELTPNPNPDIIFFDEENTGADRLMTORCKDKLNAIAISVNMOPGV 123
   61 RYEGKITRNSERKELTPNPNPDIIFFDEENTGADRLMTORCKDKLNAIAISVNMOPGV 120
DB 124 KLRVTGEMDGDGHSESLHYEGRAVDITTSDBDRSKYGMARLAVAGFDWYYESKAH 183
   121 KLRVTGEMDGDGHSESLHYEGRAVDITTSDBDRSKYGMARLAVAGFDWYYESKAH 180
QY 184 IHCSYKAENSVAAKSGCGFGSATVHLEHGCTKLVDLSPGDRVLADADGRLLYSDFLT 243
   181 IHCSYKAENSVAAKSGCGFGSATVHLEHGCTKLVDLSPGDRVLADADGRLLYSDFLT 240
QY 244 FLDRDSSKRLFYIETROPARLLTAAHLLEVAPOHNOSEATG-----STSG----- 292
   241 FLDRDSSKRLFYIETROPARLLTAAHLLEVAAP-HNDS-ATGEPEASSGSGPPSGA 298
DB 293 ---QALFASNVKPGQRYVYGE--GGQQLLPASVHSVLSREASGAAPLTAOCTILINR 347
   299 LGRPALFASVRRQRYVVAERDGRLLPAAVHSVLSREASGAAYAPLTAOCTILINR 358
QY 348 VLASCVAVIEHSHWMAHAFAPRLAOGILAAI-----CP 381
   359 VLASCVAVIEHSHWMAHAFAPRLAHLAALAPARTDRGDSGGGCGGGRVALTAP 418
DB 382 DGAIPNATTTTGIHWYSLRLYRIGSVNLDGDLHPRLGMAVAPS 425
   419 GAADAPGAGATAGIHWYSQLYQIGTWLDSALHPRLGMAVKS 462

```

RESULT 9

```

US-09-151-999-15
; Sequence 15, Application US/09151999
; Patent No. US20020151460A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Elizabeth
; TITLE OF INVENTION: REGULATION OF EPITHELIAL TISSUE BY HEDGEHOG-LIKE
; FILE REFERENCE: ONV-031.02
; CURRENT APPLICATION NUMBER: US/09/151,999
; EARLIER FILING DATE: 1998-08-11
; EARLIER APPLICATION NUMBER: 08/955,552
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 475
; TYPE: PRF
; ORGANISM: Homo sapiens Shh
; FEATURE:
; OTHER INFORMATION: Xaa at position 463 is any or unknown amino acid
US-09-151-999-15

```

```

Query Match      81.5%; Score 1807; DB 10; Length 475;
Best Local Similarity 77.8%; Pred. No. 5.3e-158;
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;

```

```

QY 4 MLLTTRLLVGFICALVSSGLTCGPGRGIGKRRHPKRLPLAYKQFIPVNAEKTIGASC 63
   1 MLLARCLLVLVSSLLVCSGLACGPGRGIGKRRHPKRLPLAYKQFIPVNAEKTIGASC 60
DB 1
QY 64 RYEGKITRNSERKELTPNPNPDIIFFDEENTGADRLMTORCKDKLNAIAISVNMOPGV 123
   61 RYEGKITRNSERKELTPNPNPDIIFFDEENTGADRLMTORCKDKLNAIAISVNMOPGV 120
DB 124 KLRVTGEMDGDGHSESLHYEGRAVDITTSDBDRSKYGMARLAVAGFDWYYESKAH 183
   121 KLRVTGEMDGDGHSESLHYEGRAVDITTSDBDRSKYGMARLAVAGFDWYYESKAH 180
QY 184 IHCSYKAENSVAAKSGCGFGSATVHLEHGCTKLVDLSPGDRVLADADGRLLYSDFLT 243
   181 IHCSYKAENSVAAKSGCGFGSATVHLEHGCTKLVDLSPGDRVLADADGRLLYSDFLT 240
QY 244 FLDRDSSKRLFYIETROPARLLTAAHLLEVAPOHNOSEATG-----STSG----- 292
   241 FLDRDSSKRLFYIETROPARLLTAAHLLEVAAP-HNDS-ATGEPEASSGSGPPSGA 298
DB 293 ---QALFASNVKPGQRYVYGE--GGQQLLPASVHSVLSREASGAAPLTAOCTILINR 347
   299 LGRPALFASVRRQRYVVAERDGRLLPAAVHSVLSREASGAAYAPLTAOCTILINR 358
QY 348 VLASCVAVIEHSHWMAHAFAPRLAOGILAAI-----CP 381
   359 VLASCVAVIEHSHWMAHAFAPRLAHLAALAPARTDRGDSGGGCGGGRVALTAP 418
DB 382 DGAIPNATTTTGIHWYSLRLYRIGSVNLDGDLHPRLGMAVAPS 425
   419 GAADAPGAGATAGIHWYSQLYQIGTWLDSALHPRLGMAVKS 462

```

RESULT 10

```

US-09-990-046-14
; Sequence 14, Application US/09990046
; Patent No. US20020156245A1
; GENERAL INFORMATION:
; APPLICANT: de Sauvage, Frederic
; TITLE OF INVENTION: Patched-2
; FILE REFERENCE: P1405R1
; CURRENT APPLICATION NUMBER: US/09/990,046
; EARLIER FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/293,505
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 437
; TYPE: PRF
; ORGANISM: Mus musculus
US-09-990-046-14

```

```

Query Match      80.7%; Score 1791; DB 9; Length 437;
Best Local Similarity 80.5%; Pred. No. 1.4e-156;
Matches 354; Conservative 21; Mismatches 43; Indels 22; Gaps 6;

```

```

QY 4 MLLTTRLLVGFICALVSSGLTCGPGRGIGKRRHPKRLPLAYKQFIPVNAEKTIGASC 63
   2 LLLARCLLVLVSSLLVCSGLACGPGRGIGKRRHPKRLPLAYKQFIPVNAEKTIGASC 61
DB 124 KLRVTGEMDGDGHSESLHYEGRAVDITTSDBDRSKYGMARLAVAGFDWYYESKAH 183
   122 KLRVTGEMDGDGHSESLHYEGRAVDITTSDBDRSKYGMARLAVAGFDWYYESKAH 181
QY 184 IHCSYKAENSVAAKSGCGFGSATVHLEHGCTKLVDLSPGDRVLADADGRLLYSDFLT 243
   182 IHCSYKAENSVAAKSGCGFGSATVHLEHGCTKLVDLSPGDRVLADADGRLLYSDFLT 241

```

```

QY 244 FLDKMDSSKLEFVITETROPBARLLTAAHLLEFVAPOHNOSEATGTSIG -OALPASVNP 302
    ||| :||:|||||: ||| ||||| ||||| ||| ||| ||||| |||
Dd 242 FLDROGAKKVFYVITELTERELLTAAHLTYAP -HNS---GPRPGSALTRSRPR 297
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 303 GORVYVVLGE--GGQQLLPASVHSVSLREASGAYAPLTAOGTILINRVLASCYAVIEHS 360
    ||||| : ||| :|||:|||||: ||| :|||:|||||: ||| :|||:|||||: |||
Dd 298 GORVYVVAERGGDRLLPAAVHSVTLREEDAGAYAPLTAAGTILINRVLASCYAVIEHS 357
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 WAHMAFAPRLAOGLLALCP-----DGAIPRA-----ATTGTGHWYSRLYRI 405
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 358 WAHRAFAFRLAHALLAALAPARTDGGGGISIPAAGSATEARGAEPTAGIHWYSOLLYHI 417
    ||| :||:||||| ||| :||:||||| ||| :||:||||| ||| :||:|||||
QY 406 GSWVLDDGDLHPLGMAVPAS 425
    ||:||||: ||||| |||
Dd 418 GTWMLDSETHMPLGMAVPAS 437
    ||:||||: ||||| |||

RESULT 11
US-09-969-520A-1
: Sequence 1, Application US/09969520A
: Patent No. US20020177163A1
: GENERAL INFORMATION:
: APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
: APPLICANT: BEACHY, Philip A.
: TITLE OF INVENTION: METHOD OF USE OF SONIC HEDEHOG PROTEIN AS A LIGAND FOR PATCHED
: FILE REFERENCE: JHU1670-1
: CURRENT APPLICATION NUMBER: US/09/969,520A
: PRIOR FILING DATE: 2002-06-04
: PRIOR APPLICATION NUMBER: US 60/235,153
: PRIOR FILING DATE: 2000-09-22
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 437
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-969-520A-1

Query Match 80.5%: Score 1791; DB 9; Length 437;
Best Local Similarity 80.5%; Pred. No. 1,4e-156;
Matches 354; Conservative 21; Mismatches 43; Indels 22; Gaps 6;

QY 4 MULTIRILVGFICALLVSGLTGCGPGRGIGKRRHPKKTPLVAKYKQFIPNVAEKTIGASG 63
    :||| :||| :||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 2 ILLLRKCFVLIVLASSLVYCGLAGCGRGGRGRHRPKKTLPLVAKQFIPNVAEKTIGASG 61
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 64 RYEGKITRNSERFKELTPVNPDIIFKDEBNTGADRLMTORCKDKLALASVNNQMPGV 123
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 62 RYEGKITRNSERFKELTPVNPDIIFKDEBNTGADRLMTORCKDKLALASVNNQMPGV 121
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 124 KLRVTEGMEDEHSHSESLHYEGRAVDITTSDBDRSKYUGMLARLAVAGPDMWYTESKAH 183
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 122 KLRVTEGMEDEHSHSESLHYEGRAVDITTSDBDRSKYUGMLARLAVAGPDMWYTESKAH 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 184 IHCYKAENSVAAKSGCGPESGATVHLEHGTGLVYKDLSGDRVLAADADGRLLYSDFLT 243
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 182 IHCYKAENSVAAKSGCGPESGATVHLEHGTGLVYKDLSGDRVLAADADGRLLYSDFLT 241
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 244 FLDKMDSSKLEFVITETROPBARLLTAAHLLEFVAPOHNOSEATGTSIG -OALPASVNP 302
    ||| :||:|||||: ||| ||||| ||||| ||| ||| ||||| |||
Dd 242 FLDROGAKKVFYVITELTERELLTAAHLTYAP -HNS---GPRPGSALTRSRPR 297
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 303 GORVYVVLGE--GGQQLLPASVHSVSLREASGAYAPLTAOGTILINRVLASCYAVIEHS 360
    ||||| : ||| :|||:|||||: ||| :|||:|||||: ||| :|||:|||||: |||
Dd 298 GORVYVVAERGGDRLLPAAVHSVTLREEDAGAYAPLTAAGTILINRVLASCYAVIEHS 357
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 WAHMAFAPRLAOGLLALCP-----DGAIPRA-----ATTGTGHWYSRLYRI 405
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 358 WAHRAFAFRLAHALLAALAPARTDGGGGISIPAAGSATEARGAEPTAGIHWYSOLLYHI 417
    ||| :||:||||| ||| :||:||||| ||| :||:||||| ||| :||:|||||
QY 406 GSWVLDDGDLHPLGMAVPAS 425
    ||:||||: ||||| |||
Dd 418 GTWMLDSETHMPLGMAVPAS 437
    ||:||||: ||||| |||

```

```

Db      418 GTWLDDSETMHPGLGMAYKSS 437

RESULT 12
US-10-013-310-1
; Sequence 1, Application US/10013310
; Publication No. US20020192216A1
; GENERAL INFORMATION:
; APPLICANT: Lamb, Jonathan Robert
; APPLICANT: Hoyne, Gerard Francis
; APPLICANT: Dallman, Margaret Jane
; TITLE OF INVENTION: Therapeutic Use
; FILE REFERENCE: 674525-2003
; CURRENT APPLICATION NUMBER: US/10/013, 310
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: PCT/GB00/02191
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: UK 9913350.6
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: UK 9921953.7
; PRIOR FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 437
; TYPE: PRT
; ORGANISM: House Mouse
US-10-013-310-1

Query Match      80.7%; Score 1791; DB 9; Length 437;
Best Local Similarity 80.5%; Pred. No. 1.4e-156;
Matches 354; Conservative 21; Mismatches 43; Indels 22; Gaps 6;

QY      4 MLLTRILLVGFICALLVSSGLTGGPGRGIGKRRHPKLLPRLAYKQFIPVNAEKTIGASG 63
Db      2 LLLARCEFLVTLASSLLVCPGLACGPGRGFGKRRHPKLLPRLAYKQFIPVNAEKTIGASG 61
QY      64 RYEKRTIRNSRFELPNNPDIIFKDEBENTGADRLMTORCKDKLNALAI SYMNQPGV 123
Db      62 RYEKRTIRNSRFELPNNPDIIFKDEBENTGADRLMTORCKDKLNALAI SYMNQPGV 121
QY      124 KLRTGEGMDEGHSEESLYHEGRAVDITTSDDRKSRYGMLARLAVBAGFDWYVYSKAH 183
Db      122 KLRTGEGMDEGHSEESLYHEGRAVDITTSDDRKSRYGMLARLAVBAGFDWYVYSKAH 181
QY      184 IHCSVKAENSVAAKSGGCFPGSATVHLEHGKTLKYDLSPGDRVLADADGRLLYSDFLT 243
Db      182 IHCSVKAENSVAAKSGGCFPGSATVHLEHGKTLKYDLSPGDRVLADADGRLLYSDFLT 241
QY      244 FLDDMDSRKLFFYIETROPRARLLTLAHLLLVYAQHNQSEATGSG-QALFASNVKP 302
Db      242 FLDDDECAKRFYIETLEPRERLLTLAHLLEVPAR-HNDS--GPTPGSALFAFSRVP 297
QY      303 GQRYVYVGE--GGQQLPASYVSHVSLREASGAYAPLTAQGTLLINRYLASCAVATEEHS 360
Db      298 GQRYVYVVAERGGDRRLRLPAVHVSYTLREEBAGAYAPLTAAGTLLINRYLASCAVATEEHS 357
QY      361 WAHMAFAFRLAAGLLAALCP-----DGAIPYA-----ATTTGIIHWSRLLYRI 405
Db      358 WAHRAFAFRLAHLAALAPARTDGGGGGSIPLAAGSATARAABETPAGIHWSQLLYRI 417
QY      406 GSWVLDDGDAIHPGLGMVAPAS 425
Db      418 GTWLDDSETMHPGLGMAYKSS 437

RESULT 13
US-09-733-634-16
; Sequence 16, Application US/09733634
; Publication No. US20030013646A1
; GENERAL INFORMATION:
; APPLICANT: Massachusetts General Hospital
; TITLE OF INVENTION: Method to stimulate Insulin production by pancreatic b-cells

```



```

: FILE REFERENCE: 17653/1240
: CURRENT APPLICATION NUMBER: US/09/733,634
: PRIORITY FILING DATE: 2000-12-08
: PRIOR APPLICATION NUMBER: US 60/170,282
: PRIOR FILING DATE: 1995-12-10
: NUMBER OF SEQ ID NOS: 32
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 16
: LENGTH: 437
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-733-634-16

```

Query Match	80.78;	Score 1791;	DB 9;	Length 437;
-------------	--------	-------------	-------	-------------

[illegible]

```

1      RESULT 14
2      US-09-021-660A-37
3      Sequence 37, Application US/09021660A
4      Patent No. US20010041668A1
5      GENERAL INFORMATION:
6      APPLICANT: Baron, M.
7      APPLICANT: Farrington, S.
8      APPLICANT: Belaussoff, M.
9      TITLE OF INVENTION: METHODS FOR MODULATING HEMATOPOIESIS AND VASCULAR
10     TITLE OF INVENTION: GROWTH
11     FILE REFERENCE: HUIP-P01-060
12     CURRENT APPLICATION NUMBER: US/09/021,660A
13     CURRENT FILING DATE: 2001-08-27
14     PRIOR APPLICATION NUMBER: 60/037,513
15     PRIOR FILING DATE: 1997-02-10
16     PRIOR APPLICATION NUMBER: 60/049,763
17     PRIOR FILING DATE: 1997-06-16
18     NUMBER OF SEQ ID NOS: 42
19     SOFTWARE: PatentIn ver. 2.1
20     SEQ ID NO 37
21     LENGTH: 437
22     TYPE: PRT
23     ORGANISM: Mus musculus

```

US-09-021-660A-37

Query Match	80.7%;	Score 1791;	DB 10;	Length 437;
Best Local Similarity	80.5%;	Pred. No. 1.4e-156;		
Matches 354;	Conservative 21;	Mismatches 43;	Indels 22;	Gaps 6

```

0Y 4 MLLTRILLVAFICALLVSSLLTGGPGIGKRRHPKTLPLAYKOFIPVNAEKTGGASC 63
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 LLLLARCLVTLASSLLVCPGLACGPGRGCKRRHPKTLPLAYKOFIPVNAEKTGGASC 61

0Y 64 RYEGKITRNSERFELPNNYNDIIFKDEENTGADRLMTOCKDKMLAIAISVNMOPGV 1233
Db 62 RYEGKITRNSERFELPNNYNDIIFKDEENTGADRLMTOCKDKMLAIAISVNMOPGV 1221

0Y 124 KLRTYEGHDEGHHSSESLHIEGAANDITTSDBRSTKGMALRLAVAGCDWYYESKAH 1833
Db 122 KLRTYEGHDEGHHSSESLHIEGAANDITTSDBRSTKGMALRLAVAGCDWYYESKAH 1811

0Y 184 IHC5VKAENSVAAKSGGCFPSATVHLEHGCTKVKLSDGDRVLADADGRLYSDFLT 2433
Db 182 IHC5VKAENSVAAKSGGCFPSATVHLEHGCTKVKLSDGDRVLADADGRLYSDFLT 2411

0Y 244 FLDRMDSRKLFPVIETROPARLLTLTAHLLEFVAPOHNOSEATG5TSG--QALFASNVKP 3022
Db 242 FLDRDEGAKFYVYIETLEPERRLTLTAHLLEFVAP--HNDS--GTPPGSALFASVPR 2977

0Y 303 GORVYVIGE--GGQOLLPASVHSVLSLEESAAGAPLTAGCTILINVLASCAVAIIEHS 3602
Db 298 GORVYVAAERCGDRRLPLAAHVSYTLREBEEGAAPLTAAGCTILINVLASCAVAIIEHS 3577

0Y 351 WAHMAFAEFRLAOGILALCP-----DGAIPTA-----ATTYTGIMHYSRLYRI 4059
Db 358 WAHRAFAEFRLAHALLALALAPARTDGGGGGIPAAQSEATARGAEPYAGIHMSQLLYHI 4177

0Y 406 GSNVYLDGALPLIGMVAPAS 425
Db 418 GTWLIDSETMHPRLGMNAKSS 437

```

RESULT 15
 US-08-900-220C-13
 ; Sequence 13, Application US/08900220C
 ; Patent No. US2002045206A1
 ; GENERAL INFORMATION:
 APPLICANT: Miao, Ningning
 ; Wang, Monica
 ; Mahanthappa, Nagesh K.
 ; Pang, Kevin
 ; Jin, Ping
 TITLE OF INVENTION: Method of Treating Dopaminergic and
 ; GABA-nergic Disorders
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 STREET: ONE POST OFFICE SQUARE
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII (text)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/900,220C
 FILING DATE: 24-Jul-1997
 CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: ONV-044.01
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-08-900-220C-13

Query Match 80.68; Score 1788; DB 8; Length 437;

Best Local Similarity 80.28; Pred. No. 2.6e-156;
Matches 353; Conservative 22; Mismatches 43; Indels 22; Gaps 6;

```
OY 4 MLTLRLVLPICALVSSGLTGPGRIGRRHRPKLTPAYKQFIPNVAEKTLAGS 63
   :||| ||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 LILLARCLVLIASSLVCPGLACGPRGFGRRHRPKLTPAYKQFIPNVAEKTLAGS 61
OY 64 RYEGKITRNSERFKELTPNYNDIIFKDENTGADRLMTORCKDLNALAISVNNQWPGV 123
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 RYEGKITRNSERFKELTPNYNDIIFKDENTGADRLMTORCKDLNALAISVNNQWPGV 121
OY 124 KLRYTEGDEGHHSESLHTEGRAVDITTSDRDSKIGMLARLAVENGFDWVYESKAH 183
   :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 RLRYTEGDEGHHSESLHTEGRAVDITTSDRDSKIGMLARLAVENGFDWVYESKAH 181
OY 184 IHCYKAENSVAAKSGGCFPGSATVHLEHGGTKLVKDLSPGDRVLAADADGRILYSDFLT 243
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 IHCYKAENSVAAKSGGCFPGSATVHLEHGGTKLVKDLSPGDRVLAADADGRILYSDFLT 241
OY 244 FLDRMDSRKLFEYIETROPARLLLTAAHLLEVAPOHNOSEATGSTG-QALFASNYKP 302
   ||| :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 FLDRDEGAKKVEYVETLEPRERLTLTAHLLEVAHP-HNDS--GPTPGPSALFASRYRP 297
OY 303 GQRYVYVGE--GGQQLPASHVSLSREASGATPPLTAOGTILINRYLASCIYVIEEHS 360
   ||||| :| | :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 298 GQRYVYVVERGDDRRLPAAVHSTLREEAGATPPLTAHGTILINRYLASCIYVIEEHS 357
OY 361 WAHNAFAPFRLAQLALACP-----DGAIPFA-----ATTGTGHHYSRLLYRI 405
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 358 WAHNAFAPFRLAHLALALAPARTDGGGGGSIIPAQASATEARGAEPAGIHHSQLLYHI 417
OY 406 GSWVIDGDALHPLGKVPAS 425
   ||| :||| :||| ||||| :|
Db 418 GTWLIDSETMHPGLGMAVKSS 437
```

Search completed: February 20, 2003, 10:14:41
Job time : 9.62902 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:08:18 ; Search time 13.9865 seconds
(without alignments)
2921.176 Million cell updates/sec

Title: US-09-827-110a-10
Perfect score: 2218
Sequence: 1 MVEMLLTRILLVGFICALL.....GSWLDGDLHPLHGVAPAS 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2218	100.0	425	2 A49424	patterning protein
2	1791	80.7	437	2 A49425	Sonic hedgehog pro
3	1780	80.3	437	2 B53193	hedgehog homolog v
4	1533	69.1	444	2 S56765	morphogen Xhh prec
5	1434	67.4	418	2 A53193	hedgehog homolog v
6	1429	64.4	415	2 A49426	sonic hedgehog gen
7	1200	54.1	396	2 B49425	Desert hedgehog pr
8	1052	47.4	336	2 C49425	Indian hedgehog pr
9	925	41.7	471	2 A46400	segment polarity p
10	449	20.2	94	2 G02735	desert hedgehog -
11	191.5	8.6	615	2 T29550	hypothetical prote
12	166	7.5	1207	2 T23754	hypothetical prote
13	158	7.1	1226	2 T24504	hypothetical prote
14	155.5	7.0	484	2 T34504	hypothetical prote
15	152.5	6.9	1021	2 T23252	hypothetical prote
16	142.5	6.4	481	2 T27665	hypothetical prote
17	130	5.9	629	2 T19563	hypothetical prote
18	128.5	5.8	868	2 T22281	hypothetical prote
19	115.5	5.2	313	2 T15855	hypothetical prote
20	114.5	5.2	481	2 T27975	hypothetical prote
21	105.5	4.8	1464	2 T13716	bazooka gene prote
22	103	4.6	205	2 T26220	hypothetical prote
23	102	4.6	477	2 G97611	UDP-N-acetyluramo
24	102	4.6	477	2 AC2834	UDP-MurNAC-pentape
25	102	4.6	503	2 C82068	aminopeptidase A/I
26	101	4.6	503	1 APECA	leucyl aminopeptid
27	101	4.6	503	2 E91283	aminopeptidase A/I
28	101	4.6	503	2 G86124	aminopeptidase A/I
29	100	4.5	503	2 AE1061	leucyl aminopeptid

30	99	4.5	811	2 F83451	probable cation-tr
31	98	4.4	1147	1 MMXIB	myosin heavy chain
32	97.5	4.4	1137	2 JC5950	integrin alpha-7 c
33	97	4.4	503	2 AC0418	leucyl aminopeptid
34	97	4.4	506	2 S55786	DNA nucleotidylexo
35	97	4.4	547	1 A23595	DNA nucleotidylexo
36	97	4.4	591	2 E83039	probable thiol-dis
37	96.5	4.4	1145	2 B75625	hypothetical prote
38	96	4.3	416	2 T02194	probable pectinase
39	95	4.3	893	2 E95053	cell wall surface
40	94.5	4.3	534	2 A32347	fimbrial protein t
41	94.5	4.3	588	2 AE2415	WD-repeat protein
42	94.5	4.3	924	2 T00518	hypothetical prote
43	94.5	4.3	1995	2 G81044	hemagglutinin/hemo
44	94.5	4.3	2124	2 T28658	polyketide synthas
45	94	4.2	481	2 G98044	UDP-N-acetyluramo

ALIGNMENTS

RESULT 1									
A49424									
patterning protein Sonic hedgehog precursor - chicken									
C:Species: Gallus gallus (chicken)									
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000									
C:Accession: A49424									
R:Riddle, R.D.; Johnson, R.L.; Laufer, E.; Tablin, C.									
A:Title: Sonic hedgehog mediates the polarizing activity of the ZPA.									
A:Reference number: A49424; MOID:94094333; PMID:8269518									
A:Accession: A49424									
A>Status: preliminary									
A:Molecule type: mRNA									
A:Residues: 1-425 <RID>									
A:Cross-references: GB:128099; NID:9453526; PIDN:AAA72428.1; PID:9453527									
A:Superfamily: sonic hedgehog protein									
F:1-26/Domain: signal sequence #status predicted <SIG>									
Query Match									
Best Local Similarity 100.0%; Pred. No. 1.8e-176;									
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MVEMLLTRILLVGFICALVSSGLTGCPRGICRRHPRKLPPLAYKQFIPNVAERTLG	60						
DB	1	MVEMLLTRILLVGFICALVSSGLTGCPRGICRRHPRKLPPLAYKQFIPNVAERTLG	60						
QY	61	ASGRYBEKTRNSERFELTPNPNPDIIFKDENTGADRLMTORCKDLNALAISVNNQW	120						
DB	61	ASGRYBEKTRNSERFELTPNPNPDIIFKDENTGADRLMTORCKDLNALAISVNNQW	120						
QY	121	PGVLRVTEGMDDECHHSSESIHGEKRAVDITTSDBRSRYGMARLAVGEFPMVYVES	180						
DB	121	PGVLRVTEGMDDECHHSSESIHGEKRAVDITTSDBRSRYGMARLAVGEFPMVYVES	180						
QY	181	KAHHCYKAKENSVAAKSGCFPSATVHLHENGTKLVKDLSPDRVLADADGRLLYSD	240						
DB	181	KAHHCYKAKENSVAAKSGCFPSATVHLHENGTKLVKDLSPDRVLADADGRLLYSD	240						
QY	241	FLTFIDMDSSRKLFYIETROPARLLLTNAHLFTVAPQHNSEATGTSGLAFASNV	300						
DB	241	FLTFIDMDSSRKLFYIETROPARLLLTNAHLFTVAPQHNSEATGTSGLAFASNV	300						
QY	301	KPGGRVVLVEGGQQLPASYHSVSLREPSGAAPLTAGCTILINNVLASCVVIEHS	360						
DB	301	KPGGRVVLVEGGQQLPASYHSVSLREPSGAAPLTAGCTILINNVLASCVVIEHS	360						
QY	361	WAHAFAFPRIAGCLALCPDGAIPTAATTTGIIHYSRLLYRIGSWLDGDLHPLGM	420						
DB	361	WAHAFAFPRIAGCLALCPDGAIPTAATTTGIIHYSRLLYRIGSWLDGDLHPLGM	420						
QY	421	VAPAS 425							
DB	421	VAPAS 425							

Db 421 VAPAS 425

RESULT 2 A49425

Sonic hedgehog protein precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999

C:Accession: A49425
R:Echelard, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; McMane
Cell 75, 1417-1430, 1993

A:Title: Sonic hedgehog, a member of a family of putative signaling molecules, is implied
A:Reference number: A49425; MUID:9409434; PMID:7916661

A:Accession: A49425
A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA
A:Residues: 1-437 <ECH>

A:Cross-references: GB:X76290

A:Note: authors translation is shown for the codon TCC at position 436

C:Genetics:
A:Gene: Shh
C:Superfamily: sonic hedgehog protein

Query Match 80.7%; Score 1791; DB 2; Length 437;

Best Local Similarity 80.5%; Pred. No. 5,9e-141;

Matches 354; Conservative 21; Mismatches 43; Indels 22; Gaps 6;

4 MLTLRLLVGFICALVSSGLTCGPGRGIGRRHKKLTPLAYKOFIPNVAEKTIGASG 63

2 LLLARCFVLIALSSLVCGGLACGPGRGGRHKKLTPLAYKOFIPNVAEKTIGASG 61

64 RREGKTRNSERKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVNMQPGV 123

62 RREGKTRNSERKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVNMQPGV 121

124 KLRTVEGMDEDHSESLHYEGRAVDITTSDRDSKYGMLARLAEAGFDWVYESKHA 183

122 KLRTVEGMDEDHSESLHYEGRAVDITTSDRDSKYGMLARLAEAGFDWVYESKHA 181

184 IHCSVKAENSVAASGCGPGSATVHLHGRTKLVKDLSPGDRVLAADGRLYSDFLT 243

182 IHCSVKAENSVAASGCGPGSATVHLHGRTKLVKDLSPGDRVLAADGRLYSDFLT 241

244 FLDRDSSKRLRYETROPARLLTAAHLFPAPQHNSEATGSGQA-LFASNVKP 302

242 FLDRDSSKRLRYETROPARLLTAAHLFPAPQHNSEATGSGQA-LFASNVKP 297

303 GORVYVIGE--GGQOLLPASVHSVLSREASGAYAPLTNOGTLINRVLASCYAVIEHS 360

298 GORVYVIAERGGDRLLPAVAHSVTLREAGAYAPLTNOGTLINRVLASCYAVIEHS 357

361 WAHMAFAPRLAAGLLALCP-----DGAIPTA-----ATTGTGHHYSRLLYRI 405

358 WAHMAFAPRLAAGLLALCP-----DGAIPTA-----ATTGTGHHYSRLLYRI 417

406 GSWVLDDGALHPLGMVAPAS 425

418 GTWLDDSETMHPDGMVAPAS 437

RESULT 3 B53193

hedgehog homolog vhh-1 - rat

C:Species: Rattus norvegicus (Norway rat)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999

C:Accession: B53193
R:Roelink, H.; Augsbury, A.; Heemskerk, J.; Kozh, V.; Norlin, S.; Ruiz i Altaba, A.;

Cell 76, 761-775, 1994

A:Title: Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedgehog
A:Reference number: B53193; MUID:94170375; PMID:8124714

A:Accession: B53193
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-437 <ROE>
A:Cross-references: GB:L27340; NID:9452122; PIDN:AAA20999.1; PID:9452123
C:Superfamily: sonic hedgehog protein

Query Match 80.3%; Score 1780; DB 2; Length 437;

Best Local Similarity 80.0%; Pred. No. 4,9e-140;

Matches 352; Conservative 22; Mismatches 44; Indels 22; Gaps 6;

4 MLTLRLLVGFICALVSSGLTCGPGRGIGRRHKKLTPLAYKOFIPNVAEKTIGASG 63

2 LLLARCFVLIALSSLVCGGLACGPGRGGRHKKLTPLAYKOFIPNVAEKTIGASG 61

64 RREGKTRNSERKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVNMQPGV 123

62 RREGKTRNSERKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVNMQPGV 121

124 KLRTVEGMDEDHSESLHYEGRAVDITTSDRDSKYGMLARLAEAGFDWVYESKHA 183

122 KLRTVEGMDEDHSESLHYEGRAVDITTSDRDSKYGMLARLAEAGFDWVYESKHA 181

184 IHCSVKAENSVAASGCGPGSATVHLHGRTKLVKDLSPGDRVLAADGRLYSDFLT 243

182 IHCSVKAENSVAASGCGPGSATVHLHGRTKLVKDLSPGDRVLAADGRLYSDFLT 241

244 FLDRDSSKRLRYETROPARLLTAAHLFPAPQHNSEATGSGQA-LFASNVKP 302

242 FLDRDSSKRLRYETROPARLLTAAHLFPAPQHNSEATGSGQA-LFASNVKP 297

303 GORVYVIGE--GGQOLLPASVHSVLSREASGAYAPLTNOGTLINRVLASCYAVIEHS 360

298 GORVYVIAERGGDRLLPAVAHSVTLREAGAYAPLTNOGTLINRVLASCYAVIEHS 357

361 WAHMAFAPRLAAGLLALCP-----DGAIPTA-----ATTGTGHHYSRLLYRI 405

358 WAHMAFAPRLAAGLLALCP-----DGAIPTA-----ATTGTGHHYSRLLYRI 417

406 GSWVLDDGALHPLGMVAPAS 425

418 GTWLDDSETMHPDGMVAPAS 437

RESULT 4 S56765

morphogen xhh precursor - African clawed frog

C:Species: Xenopus laevis (African clawed frog)
C:Date: 10-Oct-1995 #sequence_revision 03-Jan-1995 #text_change 21-Jul-2000

C:Accession: S56765
R:Stolow, M.A.; Shi, Y.B.

Nucleic Acids Res. 23, 2555-2562, 1995

A:Title: Xenopus sonic hedgehog as a potential morphogen during embryogenesis and thy

A:Reference number: S56765; MUID:95357169; PMID:7630736

A:Accession: S56765
A:Status: preliminary; nucleic acid sequence not shown

A:Residues: 1-444 <STO>
A:Cross-references: EMBL:L39213; NID:9790937; PIDN:AAC42227.1; PID:9790938

C:Superfamily: sonic hedgehog protein

Query Match 69.1%; Score 1533; DB 2; Length 444;

Best Local Similarity 66.7%; Pred. No. 1,7e-119;

Matches 301; Conservative 41; Mismatches 73; Indels 36; Gaps 5;

4 MLTLRLLVGFICALVSSGLTCGPGRGIGRRHKKLTPLAYKOFIPNVAEKTIGAS 62

1 MLVATQSLLSLFSCTIVTPPLACGPGRGIGRRHKKLTPLAYKOFIPNVAEKTIGAS 60

63 GRYEKGTRNSERKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVNMQPG 122

61 GRYEKGTRNSERKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVNMQPG 120

123 VLKRTVEGMDEDHSESLHYEGRAVDITTSDRDSKYGMLARLAEAGFDWVYESKA 182

121 VLKRTVEGMDEDHSESLHYEGRAVDITTSDRDSKYGMLARLAEAGFDWVYESKA 180

QY 183 IHCSVKAENSVAAKSGCGPFSATVHLEHGCTKLVKDLSPEDRVLAADADGRLLYSDFL 242
 DB 181 IHCSVKAENSVAAKSGCGPFSATVHLEHGCTKLVKDLSPEDRVLAADADGRLLYSDFL 240
 QY 243 TFLDMDSSRKLEFYIETFPORARLLTAAHLLFVAPOHNSOEAAGSTSGALAFASNVKP 302
 DB 241 MTDDEKDYKLEFYIETFPORARLLTAAHLLFVAPOHNSOEAAGSTSGALAFASNVKP 294
 QY 303 GORVVLGEGGQQLLPASVSVSLREASGAVAPLTAOGTILINRYLASCYAVIEEHSMA 362
 DB 295 GDLITTAAPKTYTLKAVYKEVDL-EEDTGATAPLTAHGTIVIDVYLASCYAVIEEHTMA 353
 QY 363 HMAFAPRLAQLLAALCPDGAIPTAAT-----TTTG 394
 DB 354 HMAFAPRLAQLLAALCPDGAIPTAAT-----TTTG 394
 QY 395 IHWSRLLYRIGSWVLDDGALHPLGMVAPAS 425
 DB 414 IHWSRLLYRIGSWVLDDGALHPLGMVAPAS 444

RESULT 5

A53193

hedgehog homolog vhh-1 - zebra fish

C:Species: Brachydanio rerio (zebra fish)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999

C:Accession: A53193

R:Roelink, H.; Augsburg, A.; Heemkerk, J.; Kozh, V.; Norlin, S.; Ruiz, I. Altaba, A.;

Cell 76, 761-775, 1994

A:Title: Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedgeh

A:Reference number: A53193; MUID:94170375; PMID:8124714

A:Accession: A53193

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-418 <ROE>

A:Cross-references: GB:L27585; MID:9452159; PID:9452160

C:Superfamily: sonic hedgehog protein

Query Match 67.4%; Score 1494; DB 2; Length 418;

Best Local Similarity 68.4%; Pred. No. 2.8e-116;

Matches 290; Conservative 50; Mismatches 76; Indels 8; Gaps 3;

QY 4 MLTLRILLVGFICALVSSGLTCGPRGIGRRHPRKLTPLAYKQFIPNVAEKTIGASG 63
 DB 1 MLTLRILLVLSLTLSLVSGIACGPRGIGRRHPRKLTPLAYKQFIPNVAEKTIGASG 60
 QY 64 RREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPCV 123
 DB 61 RREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPCV 120
 QY 124 KLRVTEGMDHDEHSESLHYEGRAVDITTSRDNSKCYMLARLVAEAFDWWYTESKAH 183
 DB 121 KLRVTEGMDHDEHSESLHYEGRAVDITTSRDNSKCYMLARLVAEAFDWWYTESKAH 180
 QY 184 IHCSVKAENSVAAKSGCGPFSATVHLEHGCTKLVKDLSPGDRVLAADADGRLLYSDFL 243
 DB 181 IHCSVKAENSVAAKSGCGPFSATVHLEHGCTKLVKDLSPGDRVLAADADGRLLYSDFL 240
 QY 244 FLDRDSSRKLEFYIETFPORARLLTAAHLLFVAPOHNSOEAAGSTSGALAFASNVKP 303
 DB 241 FLDRDSSRKLEFYIETFPORARLLTAAHLLFVAPOHNSOEAAGSTSGALAFASNVKP 296
 QY 304 QRVYVLGEGGQQLLPASVSVSLREASGAVAPLTAOGTILINRYLASCYAVIEEHSMAH 363
 DB 297 QRVYVLGEGGQQLLPASVSVSLREASGAVAPLTAOGTILINRYLASCYAVIEEHSMAH 354
 QY 364 WAFAPRLAQLLAALCPDGAIPTAAT--TTGIIHWSRLLYRIGSWVLDDGALHPLGMV 421
 DB 355 WAFAPRLAQLLAALCPDGAIPTAAT--TTGIIHWSRLLYRIGSWVLDDGALHPLGMV 414
 QY 422 APAS 425

DB 415 VNSS 418
 RESULT 6
 A49426
 sonic hedgehog gene shh protein - zebra fish
 N:Alternate names: local cell-cell interaction signaling protein
 C:Species: Brachydanio rerio (zebra fish)
 C:Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 10-Dec-1999
 C:Accession: A49426
 R:Krauss, S.; Concordet, J.P.; Ingham, P.W.
 Cell 75, 1431-1444, 1993
 A:Title: A functionally conserved homolog of the Drosophila segment polarity gene hh
 A:Reference number: A49426; MUID:94094355; PMID:8269519
 A:Accession: A49426
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-415 <KRA>
 A:Experimental source: embryo
 A:Note: sequence extracted from NCBI backbone (NCBIF:142459)
 C:Superfamily: sonic hedgehog protein

Query Match 64.4%; Score 1429; DB 2; Length 415;

Best Local Similarity 66.4%; Pred. No. 7e-111;

Matches 283; Conservative 49; Mismatches 72; Indels 22; Gaps 5;

QY 4 MLTLRILLVGFICALVSSGLTCGPRGIGRRHPRKLTPLAYKQFIPNVAEKTIGASG 63
 DB 1 MLTLRILLVLSLTLSLVSGIACGPRGIGRRHPRKLTPLAYKQFIPNVAEKTIGASG 60
 QY 64 RREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPCV 123
 DB 61 RREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPCV 120
 QY 124 KLRVTEGMDHDEHSESLHYEGRAVDITTSRDNSKCYMLARLVAEAFDWWYTESKAH 183
 DB 121 KLRVTEGMDHDEHSESLHYEGRAVDITTSRDNSKCYMLARLVAEAFDWWYTESKAH 180
 QY 184 IHCSVKAENSVAAKSGCGPFSATVHLEHGCTKLVKDLSPGDRVLAADADGRLLYSDFL 243
 DB 181 IHCSVKAENSVAAKSGCGPFSATVHLEHGCTKLVKDLSPGDRVLAADADGRLLYSDFL 240
 QY 244 FLDRDSSRKLEFYIETFPORARLLTAAHLLFVAPOHNSOEAAGSTSGALAFASNVKP 303
 DB 241 FLDRDSSRKLEFYIETFPORARLLTAAHLLFVAPOHNSOEAAGSTSGALAFASNVKP 296
 QY 304 QRVYVLGEGGQQLLPASVSVSLREASGAVAPLTAOGTILINRYLASCYAVIEEHSMAH 363
 DB 297 QRVYVLGEGGQQLLPASVSVSLREASGAVAPLTAOGTILINRYLASCYAVIEEHSMAH 354
 QY 364 WAFAPRLAQLLAALCPDGAIPTAATTTG--IHWYSR-----LVRIGSWVLDDGDA 414
 DB 355 WAFAPRLAQLLAALCPDGAIPTAATTTG--IHWYSR-----LVRIGSWVLDDGDA 407
 QY 415 LHPLGM 420
 DB 408 LHPLGM 413
 RESULT 7
 B49425
 Desert hedgehog protein precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999
 C:Accession: B49425
 R:Reichardt, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; M
 Cell 75, 1417-1430, 1993
 A:Title: Sonic hedgehog, a member of a family of putative signaling molecules, is imp
 A:Reference number: A49425; MUID:94094354; PMID:7916661
 A:Accession: B49425
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-396 <ECH>

A;Cross-references: GB:X:76292; NID:g443941; PIDN:CAM53924.1; PID:g443942
C:Genetics:
A:Gene: Dhh
C:Superfamily: sonic hedgehog protein

Query Match	54.1%	Score 1200;	DB 2;	Length 396;
Best Local Similarity	60.7%	Pred. No. 7.2e-92;		
Matches 241; Conservative	56;	Mismatches 84;	Indels 16;	Gaps 9;

12 LVGFIC-ALLVSSGLTCGPRG-IGKRHRPKK-LTPLAYKQFIPNVAEKT LGASGRYEK 688

Db 7 LPLCCLALLALSQSCGPGRGVGRRRYRKQLVPLLYKQFVPSMPERTLGASGPAEGR 66

```

69 ITRSRERFKELTPNNPNPDIIFRDEENGCADRLMTQRCADKLNALAIISVMNQPGVLRVT 128
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 67 VTRSRERFDLVPNPNPDIEKDEENGADRLMTGERCKERVNALAIAYMMNMGPGLRVYT 126

```

0y 129 EGMDEDHHSESLHIEGRAVDITTSRDRSKYGLARAVEAGDWMVYESKAHHCVS 188
|||||:::|||||:::|||||:::|||||:::|||||
Db 127 EGMDEDGHHNQDSLHIEGRALDITTSRDRNNKGLLARAVEAGDWMVYESRNHHSV 186
|||||:::|||||:::|||||:::|||||:::|||||

Qy 189 KAESVAASGGCGPFSATYHLEHGCTKLVKDUSPGDRVLAADADGRLLYSDFLFLDRM 248
 ||||| ::|||::||| | ::| | ||||| ||:: : | |||
 Db 187 KADSLAVRAGGCGFONATYVLRSGEKGRLRHGDWVLAADAGRVETPVLLFLDRD 246

```

0Y      249 DSSKLFVIETRÖPARLLITTAHLIFVAPQINÖSEATGSGALFASNVKPGGRVYV 308
        | | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      247 LÖRASFAVAVETERPPRKLITLPMHLVFAA--RGPAPAGDFA--PVFARLRACGS--V 300

```

```

0y 309 LGE6GQOLLPA5VSHVSLREASGAVAPLTAOITILINVLASCYAVIEEHSWAMAFAP 368
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 301 LAP6GDALDAPARVARVA - REAAVGVFAFPLTAGITLIVNDVLA5CYAVLESHQWAMHAFAP 359

```

```

Oy 369 FRLQGLLALCPDGAIPPLAATTTGTGHWSRLLYRI 405
    || | | | | | | | | | | | | | | | | | | | |
Db 360 LRLIHA-LGALLPGAV-----QPTGHNWYSRLLYRL 390

```

RESULT 8
C49425

C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 11-Jan-2000
C:Accession: C49425

A:Title: Sonic hedgehog, a member of a family of putative signaling molecules
A:Reference number: A49425; WUID:94094334; PMID:7916661
Cell 75, 1417-1430, 1993

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-336 <ECHO>

Query Match 47.48: Score 1052: DB 2: length 336:
C:Superfamily: sonic hedgehog protein

74 EREKEITPNYPNDIEKDEFEETGADRIETORCKDKINALASVNNQMPGVKIPVTEGGWDE 133
 best local similarity 80.45; Freq. NO. 1.42-75;
 Matches 215; Conservative 37; Mismatches 80; Indels 24; Gaps

[illegible]

```

D6      61 DGHSEESLHIEGRAVDITTSDRDRNKYGLLARLAVEAGFDWVYESKAHVCHSYSEHS 120
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
104 VAAECCSCCSCCSCCSCCSCCSCCSCCSCCSCCSCCSCCSCCSCCSCCSCCSCCSCCSCC

```

[illegible][illegible]

Db 181 AFQYIEIDPPRLALIPAHLLFIADNHTEPA----HFRATEASHVQPGQ--YLVYSGV 233

QY 314 QQLPASHVSYSLREASGAPRLTAQSTILINVLASCAVITEHSMANNAFAPRLAQ 373

Db 235 PGLPAPARAANS-THVALGSVAPRLTRNGTIVLEEDVVASCFNAVDHNLDAQDAFPRRL-- 291

Qy 374 GLLALCPDCAIPTAA---TTTGIIHMYSLRYRIGSWLVDGDLHP LGWVAPAS 425
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 292 -----FPLSLMGSWTPSEGVHSY PQMLYRLGRLLLEESYFNHPLMSGAGS 336

RESULT 9
A46400

N:Alternate names: hh protein
C:Species: Drosophila melanogaster
C:Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 10-Dec-1994

R:Tabata, T.; Eaton, S.; Kornberg, T.B.
Genes Dev. 6, 2635-2645, 1992

A:Accession: A46400
A:Status: preliminary
A:Molecule type: mRNA

A: Cross-references: GB:S66384; NID:g435848; PID:g435849
A: Experimental source: Oregon-R
A: Note: sequence extracted from NCBI backbone (NCBIN:138996, NCBIP:138997)

Gene 124, 183-189, 1993

A: Molecule type: mRNA
A: Residues: 1-471 <TAS>
A: Cross-references: GB:L05404

A: Note: Intron positions were determined from partial DNA sequence R: Lee, J. J.; von Kessler, D. P.; Parks, S.; Beachy, P. A. Cell 71, 33-50, 1992

A:Reference number: A43480; MUID:93008241; PMID:1394430
A:Accession: A43480
!Status: preliminary; not compared with conceptual translation
!Status: preliminary; not compared with conceptual translation

A;Residues: 1-471 <LEE>
 A;Cross-references: GB:L02793; NID:g157609; PID:g157610
 A;Note: sequence extracted from NCBI backbone (NCBIP:115418)
 A;Molecule type: DNA

C;comment: this protein is required for cell-cell communication.
C;genetics:
A;gene: hh
;Cross-references: F|vBase:FBam0004644

A;inclusions: 150/3; 240/1
C;Superfamily: sonic hedgehog protein
C;Keywords: transmembrane protein
62-82/Domains: transmembrane &status predictedTM

Query Match	41.78;	Score 925;	DB 2;	Length 471;
Best Local Similarity	47.88;	Pred. No. 6.8e-69;		
Matches 197. Conservative	65.	Mismatches 132.	Indels 19.	Cons

6 LTRILLVGFICALVSSGLTCGPRGICGRHRHPKLTPLAYKQFIIPNAEKTLAGSGRY 65

QY 66 EGI TRN S R F K E L T P N Y N P D I I F K D E N T G A D R L M T Q R C K D K L N A L A S Y M N Q W G V K L 125

QY 126 RYTEGWDDEGHHSERSLHYEGRAVDITTSRDRSKYGMRLAYEAGFDWYYESKAH1H 185

```

Oy      186 CSVAKEENSVAANKSGCGFCGSAATNHEHGTAKVKRGLADAGRLTYSDFFL 245
Db      243 CSVASDSSISSHVHCFTPEESTALLSEGVKRPCLGETSIDGRVLSMTANGAVYSEVLFM 302
Oy      246 DRMDSRRLLFYETTRQPPARILLTAHLLEFAPQHNSSEATGSTSGALFASNVKPGOR 305
Db      303 DRIELEOMONFOVLIHT-DEGAVALTVTPRAHLVSVMQESOKLT-----FVFADRIEKKNO 354
Oy      306 VYVLOGEGGOQLLPAPVHSV-SUREASGAYAPLTAOGTILINRVLASCAVAIEHSMAHW 364
Db      355 VLVRDETVEGELRPQRHVXGVSR--SKGVAPALETGRTIVNSVAASCVAIVNSOSLAHW 412
Oy      365 AFAFRLAAGLLA-----ALCPDAFIPTAATTGGTHMYRSRLYRIGSWVL 410
Db      413 GLAPRRLLSTLEAWLPAKEQLHSPKRYVSSAQOQGCHHYANALKKVADYVL 464

RESULT 10
G02735
desert hedgehog - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 11-Jan-2000
C:Accession: G02735
R:Drummond, I.A.
submitted to the EMBL Data Library, June 1996
A:Reference number: H01643
A:Accession: G02735
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-94 <DRU>
A:Cross-references: EMBL:U59748; NID:g1401271; PID:g1401272
C:Genetics:
A:Gene: hDHH
C:Superfamily: sonic hedgehog protein

Query Match          20.2%; Score 449; DB 2; Length 94;
Best Local Similarity 84.0%; Pred.No. 2.7e-30;
Matches 79; Conservative 13; Mismatches 2; Indels 0; Gaps 0.

Oy      87 IIFKRENTYGADRMLTORCKDKLNALSIYNQMVGVKIRVTEGDNDGHSEESLHYEG 146
Db      1 IIFKEENSGADRMLTERCKERVNALATAVMNMWPGVRLRVTEGDNDGHNAQDSLHYEG 60

Oy      147 RAVDTITSDRRSKYGMALARLAVEAGFDVVYES 180
Db      61 RAIDLTTSDRRKNKYGLLARLAVEAGFDVVYGS 94

RESULT 11
T29550
hypothetical protein ZK377.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T29550
R:Nhan, M.; Hawkins, J.
submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans coamid ZK377.
A:Reference number: Z20639
A:Accession: T29550
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-615 <NHA>
A:Cross-references: EMBL:U08183; PIDN:AA852656.1; GSPDB:GN00028; CESP:ZK377.1
A:Experimental source: strain Bristol NZ; clone ZK377
C:Genetics:
A:Gene: CESP:ZK377.1
A:Map position: X
A:introns: 46/1; 70/3; 157/2; 176/1; 245/2; 314/3; 402/3; 499/2

Query Match          8.6%; Score 191.5; DB 2; Length 615;
Best Local Similarity 28.9%; Pred.No. 9.9e-08;
Matches 58; Conservative 39; Mismatches 87; Indels 17; Gaps 7
```

```

OY 201 CFPSSATVHLEHGGFKLKLDSPGGRVLAADADRLLYSDELFTFLDRDMSRKLFVYET 260
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 416 CFSGDMEVETE-DGIKMKTKDKIDKVLSPD-EAFVITYSPVIMFLHKHDEETIAEFNILET 473
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 261 ROPRARLLTLTAHLLFVAPQHNOSSEATGSTGQALFASNVKPGORVYVLGEGGOOLLPAS 320
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 474 ANGHS-IKLTDNHLLIYVSDCSTRSDL-----KLVAKEVMKDCIHHTTD-SNVVTKKK 525
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 321 VHSVSLREASGATAPLTAOCTILINRYLASCVAIVIEHSHAHNAFAFRLAOGLLAFL- 379
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 526 VSKISKVYE-TGIYSPLSTGDIIVNRVLASCHSNLAKLSIQOTFFSLYKRTSSVFHNL 584
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 380 -----CPGPAIFPAATTTGT 395
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 585 FPKSTEEGCDLPVGEVILTSTV 605
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
T23754
hypothetical protein T05C12.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T23754; T24513
R:Thomas, K.
submitted to the EMBL Data Library, June 1995
A:Reference number: Z19793
A:Accession: T23754
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-1207 <M12>
A:Molecule type: DNA
A:Cross-references: EMBL:Z49568; PIDN:CAA90265.1; GSPDB:GN00020; CESP:T05C12.10
A:Experimental source: clone M10
R:Buton, J.
submitted to the EMBL Data Library, October 1995
A:Reference number: Z19901
A:Accession: T24513
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1207 <M12>
A:Cross-references: EMBL:Z66500; PIDN:CAA91313.1; GSPDB:GN00020; CESP:T05C12.10
A:Experimental source: clone T05C12
C:Genetics:
A:Gene: CESP:T05C12.10
A:Map position: 2
A:Introns: 31/3; 87/2; 141/3; 180/2; 203/3; 267/1; 776/2; 794/2; 834/2; 1086/3; 1143/3

Query Match 7.5%; Score 166; DB 2; Length 1207;
Best Local Similarity 29.3%; Pred. No. 3.4e-05;
Matches 55; Conservative 33; Mismatches 84; Indels 16; Gaps 6;

OY 187 SYKAANSVAASGCG-----CFPGSATVHLEHGGFKRLVADLSPGGRVLAADADGRLLYSDF 241
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 942 AVLADTPGAGAGGAGGSRNCFSPADSLV-TYVTGQRMDELDIGDVLVLPSSGNWLKYEKV 1000
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 242 LTFELDRDMSRKLFVIEETROPFARLLTLTAHLLFVAPQHNOSSEATGSTG-----QAL 295
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1001 EMEFYHREKTRTNFVYLTKSGR-KLSLTGRLHLPLVACSGVEQTTNMPDGLDVAAMRSK 1059
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 296 FASNVKPGORVYVLGEGGOOLLPASVHVSLSREASGATAPLTAOCTILINRYLASCYAV 355
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1060 YAEKARKECVLSIDSEVLADEIVRVG-RMTWGIYSPTVEGSLIVDGLVSCPSH 1117
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 356 IEESHMAH 363
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1118 LESH-AH 1124
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
T24045
hypothetical protein R08B4.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

```

C:Accession: T24045
 R:White, S.
 Submitted to the EMBL Data Library, November 1995
 A:Reference number: Z19834
 A:Accession: T24045
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1226 <WTL>
 A:Cross-references: EMBL:Z68008; PIDN:CAA92000.1; GSPDB:GN00028; CESP:R08B4.1
 A:Experimental source: clone R08B4
 C:Genetics:
 A:Gene: CESP:R08B4.1
 A:Map position: X
 A:Introns: 53/2; 113/1; 152/3; 204/3; 272/1; 354/1; 389/3; 714/3; 839/3; 877/1; 960/3; 1

Query Match 7.1%; Score 158; DB 2; Length 1226;
 Best Local Similarity 27.8%; Pred. No. 0.00016;
 Matches 52; Conservative 37; Mismatches 82; Indels 16; Gaps 7;

OY 192 NSVAAKSGCGPFGSATVHLEHGKTLVNDLSPGDRVLAADADGRLYSDFLTDRMDS 251
 DB 1012 SALVATGACFSLDMVW-TTPGKKRMDQIDIGDYVLTADLE-KTYFPPTLMIHREPEK 1069
 OY 253 RLFVIEROPRRALLLTAAHLF---VAPQHNSEATGSGALFASNKPGQRYVY 308
 DB 1070 VQEFLLTMYEYKT-LRTISRHFMYRNKCGKSPQYIKMLPHDGAIRFASLDEVDCCVV 1128
 OY 309 L--GCGGQOLLPASVHVSLEREASGAYAPLTAOCTILINRYLASCYAVIEHWSA---H 363
 DB 1129 LYKGYRQOKIETITRSV-----RTGIYSPLTNNGRITIVNDLASCYSYSLQNTQTTF 1183
 OY 364 WAFAPFR 370
 DB 1184 WAYDKLR 1190

RESULT 14

T24504
 hypothetical protein ZK1290.12 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T34504

R:Atch A.
 Submitted to the EMBL Data Library, July 1995
 A:Description: The sequence of C. elegans cosmid ZK1290.
 A:Reference number: Z21535
 A:Accession: T34504
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-484 <TAAT>
 A:Cross-references: EMBL:U21308; PIDN:AA893321.1; GSPDB:GN00020; CESP:ZK1290.12
 A:Experimental source: strain Bristol N2; clone ZK1290
 C:Genetics:
 A:Gene: CESP:ZK1290.12
 A:Map position: 2
 A:Introns: 66/3; 124/2; 181/2; 392/1

Query Match 7.0%; Score 155.5; DB 2; Length 484;
 Best Local Similarity 24.9%; Pred. No. 6.9e-05;
 Matches 50; Conservative 50; Mismatches 82; Indels 19; Gaps 7;

OY 194 VAAKSGCGPFGSATVHLEHGKTLVNDLSPGDRVLAADADGRLYSDFLTDRMDS 253
 DB 283 VASGVPACFTGNSKV-WTPAGEKSMADSLVGDVMTYEV-GKMTYTRVASWLRPLDTKA 340
 OY 254 LEVYETROPRRALLLTAAHLFVAPQHNSEATGSGALFASNKPGQRYVYLGEG 312
 DB 341 AFIKLTED-----GAIIDMPQHFYKANCYTEEMELVYVADMTIGDCLAV--KE 389
 OY 313 GQOLLPASVHVSLEREASGAYAPLTAOCTILINRYLASCYAVIE---EHSWAFAP 368
 DB 390 NEKLVTITSEKSTFE--TGYYAPWTEEGDLIVDVYASCHNVYKANTLSHTELFATSV 448

OY 369 FRLAQLAALCPDGAIPPA 389
 DB 449 OOKMRSVLGSLEBTEGHLPTS 469

RESULT 15

T23252
 hypothetical protein K02E2.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T23252

R:Lloyd, C.
 Submitted to the EMBL Data Library, November 1996
 A:Reference number: Z19716
 A:Accession: T23252
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1021 <WTL>
 A:Cross-references: EMBL:Z81560; PIDN:CAB04547.1; GSPDB:GN00023; CESP:K02E2.2
 A:Experimental source: clone K02E2
 C:Genetics:
 A:Gene: CESP:K02E2.2
 A:Map position: 5
 A:Introns: 41/1; 79/3; 131/3; 197/1; 262/1; 297/3; 566/3; 610/3; 735/3; 775/3; 834/3;

Query Match 6.9%; Score 152.5; DB 2; Length 1021;
 Best Local Similarity 27.7%; Pred. No. 0.00036;
 Matches 49; Conservative 31; Mismatches 72; Indels 25; Gaps 7;

OY 214 GTRKYKDLSPGDRVLAADADGRLYSDFLTDRMDSKLFYETROPRRALLLTAAH 273
 DB 814 GKRRDEIEIGDYVLTADLK-TALFSATLWTHREPEVQEFLEIKTNGKT-LQLTAGH 871
 OY 274 LLEFVA-----POHNOSEATGS-----TSGOALFASNVKPGQRYVYLGEGGQOL 316
 DB 872 FYIATECRILPEKNSSLNSTERYRLIDTLPDSEFKLASQIKIGECCLI--HNGDOF 929
 OY 317 LPASVHVSLEREASGAYAPLTAOCTILINRYLASCYAVIEH---SWAHAFAPFR 370
 DB 930 RMQKIDISL-KTVSTGIYSPLTENGRILIVNDLASCYSVQOQNVLTTFEFNAFDRLR 985

Search completed: February 20, 2003, 10:13:30
 Job time : 17.9865 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:07:32 ; Search time 13.0329 Seconds
(without alignments)
1352.533 Million cell updates/sec

Title: US-09-827-110a-10

Perfect score: 2218
Sequence: 1 MVEMLLRRLVGFICALL.....GSWVLDGDALHPLGVNAPAS 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2218	100.0	425	1 SHH_CHICK	091035 gallus gall
2	1807	81.5	462	1 SHH_HUMAN	015465 homo sapien
3	1791	80.7	437	1 SHH_MOUSE	062226 mus musculu
4	1780	80.3	437	1 SHH_RAT	063673 ratus norv
5	1603.5	72.3	432	1 SHH_CYNPY	090385 cynops pyr
6	1533	69.1	444	1 SHH_XENLA	092000 xenopus lae
7	1494	67.4	418	1 SHH_BRARE	092008 brachydantio
8	1458	65.7	416	1 TWHH_BRARE	090419 brachydantio
9	1295.5	58.4	408	1 IHH_CHICK	098938 gallus gall
10	1286	58.0	409	1 IHH_CHICK	091612 xenopus lae
11	1277.5	57.6	411	1 IHH_MOUSE	097812 mus musculu
12	1265.5	57.1	411	1 IHH_HUMAN	014623 homo sapien
13	1232.5	55.6	412	1 IHH_BRARE	098862 brachydantio
14	1200	54.1	396	1 DHH_MOUSE	061468 mus musculu
15	1198	54.0	396	1 DHH_HUMAN	043323 homo sapien
16	1129.5	50.9	366	1 DHH1_XENLA	091610 xenopus lae
17	1119.5	50.5	398	1 DHH2_XENLA	091611 xenopus lae
18	925	41.7	471	1 HH_DROME	002936 drosophila
19	876.5	39.5	481	1 HH_DROME	056674 drosophila
20	603	27.2	121	1 SHH_RASEL	098938 gallus ele
21	600	27.1	121	1 SHH_CARAU	079681 carassius a
22	600	27.1	121	1 SHH_PUNTE	079850 punctus tet
23	599	27.0	121	1 SHH_RASHE	079864 rasbora het
24	599	27.0	121	1 SHH_RASPA	079869 rasbora pav
25	598	27.0	121	1 SHH_TANAL	079915 tanichthys
26	597	26.9	121	1 SHH_AMBCH	079662 amblypharyn
27	595	26.8	121	1 SHH_DANAE	013235 danio aff.
28	595	26.8	121	1 SHH_DANAE	013235 danio aff.
29	595	26.8	121	1 SHH_DANAT	013238 danio aff.
30	595	26.8	121	1 SHH_DANAT	013245 danio frank
31	595	26.8	121	1 SHH_DANKE	079709 danio kerri
32	595	26.8	121	1 SHH_DANKE	079717 danio pulch
33	595	26.8	121	1 SHH_DEYDE	013241 devario dev

34	595	26.8	121	1 SHH_DEYMA	013247 devario mal
35	595	26.8	121	1 SHH_DEYPA	013250 devario pat
36	591	26.6	121	1 SHH_PUNCO	079830 punctus con
37	396	17.9	88	1 DHH_BRARE	079729 brachydantio
38	293	13.2	58	1 IHH_CARAU	079693 carassius a
39	293	13.2	58	1 IHH_DANAT	013240 danio aff.
40	293	13.2	58	1 IHH_DANKE	079711 danio kerri
41	293	13.2	58	1 IHH_DANPU	079719 danio pulch
42	293	13.2	58	1 IHH_PUNTE	013243 devario dev
43	293	13.2	58	1 IHH_PUNTE	079852 punctus tet
44	293	13.2	58	1 IHH_RASEL	079860 rasbora ele
45	289	13.0	58	1 SHH_PSEPR	079839 pseudorasbo

ALIGNMENTS

RESULT 1	ID	SHH_CHICK	STANDARD;	PRT;	425 AA.
AC	091035;				
DT	15-JUL-1999 (Rel. 38, Created)				
DT	15-JUL-1999 (Rel. 38, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Sonic hedgehog protein precursor (SHH).				
GN	SHH.				
OS	Gallus gallus (Chicken).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;				
OC	Gallus.				
ON	NCBI_TaxID=9031;				
OX	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Limb bud;				
RX	MEDLINE=94094333; PubMed=8269518;				
RA	Riddle R.D., Johnson R.L., Lauffer E., Tablin C.;				
RT	"Sonic hedgehog mediates the polarizing activity of the ZPA.";				
RL	Cell 75:1401-1416(1993).				
RN	[2]				
RP	FUNCTION. AND AUTOPROTEOLYTIC CLEAVAGE.				
RX	MEDLINE=95254654; PubMed=7736596;				
RA	Koelink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,				
RT	Beachy P.A., Jessell T.M.;				
RT	"Floor plate and motor neuron induction by different concentrations of the amino-terminal cleavage product of sonic hedgehog autoproteolysis.";				
RL	Cell 81:445-455(1995).				
CC	-I- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION.				
CC	-I- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).				
CC	-I- TISSUE SPECIFICITY: EXPRESSED IN THE POSTERIOR LIMB BUD MESENCHYME, THE HENSEN'S NODE, THE NOTOCHORD, AND THE FLOOR PLATE OF THE NEURAL TUBE.				
CC	-I- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT STAGE 17 DURING THE INITIATION OF LIMB BUD FORMATION. FROM THAT POINT ONWARDS, THE EXPRESSION PATTERN EXACTLY MATCHES THE LOCATION OF THE ZONE OF POLARIZING ACTIVITY (ZPA).				
CC	-I- INDUCTION: BY RETINOIC ACID.				
CC	-I- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN				

CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: L28099; AAA72428.1; -
 DR HSP: 062226; 1VHH.
 DR MEROPS: C46.001; -
 DR InterPro: IPR000320; HH_signal.
 DR InterPro: IPR001767; Hedgehog_hntc.
 DR InterPro: IPR003586; Hedgehog_hntc.
 DR InterPro: IPR003587; Hedgehog_hntc.
 DR InterPro: IPR002203; Interh.
 DR Pfam: PF01079; Hntc; 1.
 DR Pfam: PF01085; HH_signal; 1.
 DR PRINTS: PR00632; SONICHHOG.
 DR PRODOM: PD003042; HH_signal; 1.
 DR SMART: SM00305; Hntc; 1.
 DR SMART: SM00306; Hntc; 1.
 DR PROSITE: PS00817; INTER_N_TER; 1.
 DR Developmental protein: Autocatalytic cleavage; Hydrolyase; Protease;
 DR Signal: Lipoprotein; Palmitate.
 FT SIGNAL 1 26
 FT CHAIN 27 425
 FT CHAIN 27 200
 FT CHAIN 201 425
 FT SITE 200 201
 FT SITE 246 246
 FT SITE 270 270
 FT ACT_SITE 273 273
 FT BINDING 200 200
 FT DOMAIN 390 393
 FT LIPID 27 27
 SQ SEQUENCE 425 AA: 46474 MW: 5496274430440173 CRC64:
 Query Match 100.0%; Score 2218; DB 1; Length 425;
 Best Local Similarity 100.0%; Pred. No. 1.7e-172;
 Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 241 FLTLDRMDSRKFLFYETTRPPARLLTLTAHLFLVAPQHNSERTGTSQALFASNV 300
 QY 301 KGORRYVYEGEGGOQLLPASVSVSLREBAGAVAPLTAGTILLNRVASCYAVIEHS 360
 DB 301 KGORRYVYEGEGGOQLLPASVSVSLREBAGAVAPLTAGTILLNRVASCYAVIEHS 360
 QY 361 WAHMAFAPRLAOGLLAALCPDGAIPTAATTTGTHWSRLRYRGSVWLDGDLHPGLM 420
 DB 361 WAHMAFAPRLAOGLLAALCPDGAIPTAATTTGTHWSRLRYRGSVWLDGDLHPGLM 420
 QY 421 VAPAS 425
 DB 421 VAPAS 425
 RESULT 2
 SHH_HUMAN STANDARD; PRT; 462 AA.
 ID SHH_HUMAN
 AC Q15465;
 DT 15-JUL-1999 (Rel. 38, Created)
 DR 15-JUL-1999 (Rel. 38, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 GN Sonic hedgehog protein precursor (SHH) (HHC-1).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal Lung;
 RX MEDLINE=96070431; PubMed=7590746;
 RA Marigo V., Roberts D.J., Lee S.M.K., Tsukurov O., Levi T.,
 RA Gastier J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seldman C.E.,
 RA Jenkins N.A., Seldman J.G., McMahon A.P., Tabin C.;
 RT "Cloning, expression, and chromosomal location of SHH and HH: two
 RT human homologues of the Drosophila segment polarity gene hedgehog";
 RL Genomics 28:44-51(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC Tate G., Kishimoto K., Mitsuya T.;
 RT "Expression of Sonic hedgehog and its receptor Patched/Smoothed in
 RT human cancer cell lines and embryonic organs";
 RL J. Biochem. Mol. Biol. Biophys. 4:27-34(2000).
 RN [3]
 RP SEQUENCE OF 1-187 FROM N.A.
 RA Strong C., Graves T., Sutterer C., Ozerisky P.;
 RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 119-167 FROM N.A.
 RX MEDLINE=95236997; PubMed=7720571;
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
 RA Zhao R., Seldman M.F., Fallon J.F., Beachy P.A.;
 RT "Products, genetic linkage and limb patterning activity of a murine
 RT hedgehog gene";
 RL Development 120:3339-3353(1994).
 RN [5]
 RP PALMITOYLATION.
 RX MEDLINE=98256339; PubMed=9593755;
 RA Replinsky R.B., Zeng C., Wen D., Rayhorn P., Baker D.P., Williams K.P.,
 RA Bixler S.A., Ambrose C.M., Garber E.A., Mlatkowski K., Taylor F.R.,
 RA Wang E.A., Galdes A.;
 RT "Identification of a palmitic acid-modified form of human Sonic
 RT hedgehog";
 RL J. Biol. Chem. 273:14037-14045(1998).
 RN [6]
 RP VARIANTS HPE ARG-31; GLY-117 AND ARG-117.
 RX MEDLINE=97051937; PubMed=8896572;
 RA Roessler E., Belloni E., Gaudentz K., Jay P., Berta P., Scherer S.W.,
 RA Tsui L.-C., Muenke M.;
 RT "Mutations in the human Sonic hedgehog gene cause holoprosencephaly";
 RL Nat. Genet. 14:357-360(1996).
 RN [7]

RP VARIANTS HPE ARG-31; GLY-117; ARG-117; GLU-224; THR-226 AND THR-383.
 RX MEDLINE=96027036; PubMed=9302262;
 RA Roessler E., Belloni E., Gaudenz K., Vargas F., Scherer S.W.,
 RA Tsai L.-C., Muenke M.;
 RT "Mutations in the C-terminal domain of Sonic hedgehog cause
 RT holoprosencephaly";
 RL Hum. Mol. Genet. 6:1847-1853(1997).
 RN [8]
 RP VARIANTS HPE HIS-100; GLN-188 AND ASN-222.
 RX MEDLINE=99371775; PubMed=10441331;
 RA Odent S., Attli-Bitach T., Blayau M., Mathieu M., Aug J.,
 RA Delezo de A.L., Gall J.Y., Le Marec B., Munnich A., David V.,
 RA Vekemans M.;
 RT "Expression of the Sonic hedgehog (SHH) gene during early human
 RT development and phenotypic expression of new mutations causing
 RT holoprosencephaly";
 RL Hum. Mol. Genet. 8:1683-1689(1999).
 RN [9]
 RP VARIANTS HPE V-88; K-115; R-236; 263-R-A-269 DEL; D-290; A-424 AND
 RP L-436.
 RX MEDLINE=20025757; PubMed=10556296;
 RA Nanni L., Ming J.E., Bocian M., Steinhilber K., Bianchi D.W.,
 RA Die-Smulders C., Giannotti A., Imaizumi K., Jones K.L., Campo M.D.,
 RA Martin R.A., Melnick P., Pierpont M.E.M., Robin N.H., Young I.D.,
 RA Roessler E., Muenke M.;
 RT "The mutational spectrum of the sonic hedgehog gene in
 RT holoprosencephaly: SHH mutations cause a significant proportion of
 RT autosomal dominant holoprosencephaly";
 RL Hum. Mol. Genet. 8:2479-2488(1999).
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
 CC ASSOCIATION WITH SMOOTHED (SMO), TO ACTIVATE THE TRANSCRIPTION
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
 CC BY THE NOTOCHORD THAT INDICES VENTRAL CELL FATE IN THE NEURAL TUBE
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
 CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN FETAL INTESTINE, LIVER, LUNG, AND
 CC KIDNEY. NOT EXPRESSED IN ADULT TISSUES.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- DISEASE: DEFECTS IN SHH ARE THE CAUSE OF HOLOPROSENCEPHALY (HPE).
 CC HPE IS A CLINICALLY VARIABLE AND GENETICALLY HETEROGENEOUS
 CC MALFORMATION IN WHICH THE DEVELOPING FOREBRAIN FAILS TO CORRECTLY
 CC SEPARATE INTO RIGHT AND LEFT HEMISPHERES. IN THE MOST SEVERE FORM
 CC (ALOBAR HPE), THERE IS NO INTERHEMISPHERIC FISSURE, A SINGLE BRAIN
 CC VENTRICLE IS PRESENT. HPE IS ASSOCIATED WITH SEVERAL DISTINCT
 CC FACIES AND PHENOTYPIC VARIABILITY. IN THE MOST EXTREME CASES,
 CC ANOPHTHALMIA OR CYCLOPIA IS EVIDENT ALONG WITH A CONCENTRAL
 CC ABSENCE OF THE MATURE NOSE. THE LESS SEVERE FORM FEATURES FACIAL
 CC DYSMORPHIA CHARACTERIZED BY OCULAR HYPERTELORISM, DEFECTS OF THE
 CC UPPER LIP AND/OR NOSE, AND ABSENCE OF THE OLFACTORY NERVES OR
 CC CORPUS CALLOSUM. THE MAJORITY OF HPE CASES ARE APPARENTLY
 CC SPORADIC, ALTHOUGH CLEAR EXAMPLES OF AUTOSOMAL DOMINANT (AD)
 CC INHERITANCE HAVE BEEN DESCRIBED. INTERESTINGLY, UP TO 30% OF
 CC OBLIGATE CARRIERS OF HPE GENE IN AD PEDIGREES ARE CLINICALLY
 CC UNAFFECTED.

CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC CC DATABASE: NAME=Altschul Genet. Cytogenet. Oncol. Haematol.;
 CC WWW="http://www.infobio.gen.fr/services/chromocancer/Genes/SHHID378.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch)
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L36518; AAA62179.1; -;
 CC DR EMBL: AC002484; AAB67604.1; -;
 CC DR HSSP: Q62226; 1VH.
 CC DR MEROPS: C46.002; -;
 CC DR Genew: HGNC:10848; SHH.
 CC DR MIM: 600725; -;
 CC DR MIM: 142945; -;
 CC DR InterPro: IPR000320; HH_signal.
 CC DR InterPro: IPR001767; Hedgehog_hntc.
 CC DR InterPro: IPR003586; Hedgehog_hntc.
 CC DR InterPro: IPR003587; Hedgehog_hntn.
 CC DR InterPro: IPR002203; Inteln.
 CC DR InterPro: IPR001657; SonicHH.
 CC DR Pfam: PF01079; Hntc; 1.
 CC DR Pfam: PF01085; HH_signal; 1.
 CC DR PRINTS: PD0632; SONICHHOG.
 CC DR PRODOM: PD003042; HH_signal; 1.
 CC DR SMART: SM00305; Hntc; 1.
 CC DR SMART: SM00306; Hntn; 1.
 CC DR PROSITE: PS50817; INTER_N_TER; 1.
 CC DR Developmental protein: Autocatalytic cleavage; Hydrolyase; Protease;
 CC KX Signal: Iiprotein; Palmitate; Disease mutation; Holoprosencephaly.
 CC FT SIGNAL 1 23
 CC FT CHAIN 24 462
 CC FT CHAIN 24 197
 CC FT CHAIN 198 462
 CC FT SITE 197 198
 CC FT SITE 243 243
 CC FT SITE 267 267
 CC FT ACT_SITE 270 270
 CC FT BINDING 197 197
 CC FT DOMAIN 407 411
 CC FT LIPID 24 24
 CC FT LIPID 31 31
 CC FT VARIANT 88 88
 CC FT VARIANT 100 100
 CC FT VARIANT 115 115
 CC FT VARIANT 117 117
 CC FT VARIANT 117 117
 CC FT VARIANT 117 117
 CC FT VARIANT 188 188
 CC FT VARIANT 222 222
 CC FT VARIANT 224 224
 CC FT VARIANT 226 226
 CC FT VARIANT 236 236
 CC FT VARIANT 263 269
 CC Query Match 81.5%; Score 1807; DB 1; Length 462;

Best Local Similarity 77.8%; Pred. No. 4,4e-139;
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;

QY 4 MLTTRILVGCIFCALIVSSGLTGPCRGIGKRRHPPKLLPLAYKOPIPVNAEKTLCASG 63
1 MLTTRILVGCIFCALIVSSGLTGPCRGIGKRRHPPKLLPLAYKOPIPVNAEKTLCASG 60

QY 64 RYEGKTRNSEREREKELTPNTPDIIFFDEENTGADRLMTCRKDKLALAIYVNMOPGV 123
61 RYEGKTRNSEREREKELTPNTPDIIFFDEENTGADRLMTCRKDKLALAIYVNMOPGV 120

QY 124 KLRTTEGMDSDGHHSESLYEGRAVDITSDRNRKYGMRLARAVAGDMWYEEKAH 183
121 KLRTTEGMDSDGHHSESLYEGRAVDITSDRNRKYGMRLARAVAGDMWYEEKAH 180

QY 184 IHCYKAKNSVAAKSGGCGFSATVHLHSGTKLVKDLSPGDRVLADADGRLLYSDFLT 243
181 IHCYKAKNSVAAKSGGCGFSATVHLHSGTKLVKDLSPGDRVLADADGRLLYSDFLT 240

QY 244 FLDRMDSKRLFYETFRORARILLTAHLLFYAPQHNSGATG-----STSG----- 292
241 FLDRMDSKRLFYETFRORARILLTAHLLFYAPQHNSGATG-----STSG----- 298

QY 293 ---GALFASVKKPGORVYVGE--GGQQLPASVHSVSLREASGAPLTAQSTILLNR 347
299 LGRRALFASVKKPGORVYVGE--GGQQLPASVHSVSLREASGAPLTAQSTILLNR 358

QY 348 VLASCVAVIEHSHVMAHAFAPFLAOGILAL-----CP 381
359 VLASCVAVIEHSHVMAHAFAPFLAOGILAL-----CP 381

QY 382 DGAIPRTAATTTTGHWSKLLRYIGSVNLDGALHPLGMAVAPS 425
419 DGAIPRTAATTTTGHWSKLLRYIGSVNLDGALHPLGMAVAPS 462

DB 419 GAADAPGAGATAGIHWSQLLYQIGTWLDSALHPLGMAVAPS 462

RESULT 3
SHH_MOUSE STANDARD; PTG; 437 AA.
ID SHH_MOUSE
AC 062226;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sonic hedgehog protein precursor (SHH) (HMG-1).
GN SHH OR HHG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RA MEDLINE=94094334; PubMed=7916661;
RA Echeverri Y., Epstein D.J., St Jacques B., Shen L., Mohler J., McMahon J.A., McMahon A.P.;
RT "Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity."; Cell 75:1417-1430(1993).
RL (2)
RP REVISION TO 122.
RC STRAIN=C57BL/6J;
RA McMahon A.P.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN (3)
RP SEQUENCE FROM N.A. AND AUTOPROTEOLYTIC CLEAVAGE.
RX MEDLINE=95236997; PubMed=7720571;
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K., Zhao R., Seidman M.F., Fallon J.F., Beachy P.A.;
RT "Products, genetic linkage and limb patterning activity of a murine hedgehog gene."; Development 120:3339-3353(1994).
RL (4)
RP FUNCTION, AND AUTOPROTEOLYTIC CLEAVAGE.

RX MEDLINE=95236997; PubMed=7720571;
RA Roelink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T., Beachy P.A., Jessell T.M.;
RT "Floor plate and motor neuron induction by different concentrations of the amino-terminal cleavage product of sonic hedgehog autoproducts."; Cell 81:445-455(1995).
RL (5)
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 34-195.
RX MEDLINE=96069744; PubMed=7477329;
RA Hall T.M.T., Porter J.A., Beachy P.A., Leahy D.J.;
RT "A potential catalytic site revealed by the 1.7-A crystal structure of the amino-terminal signalling domain of Sonic hedgehog."; Nature 378:212-216(1995).
RL (6)
CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH FLOOR PLATE-AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN A NUMBER OF EMBRYONIC TISSUES INCLUDING THE NOTOCHORD, VENTRAL NEURAL TUBE, FLOOR PLATE, LUNG BUD, ZONE OF POLARIZING ACTIVITY AND POSTERIOR DYSPL MESCENCHYME OF LIMBS. IN THE ADULT, EXPRESSED IN LUNG AND NEURAL RETINA.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE DURING GASTRULATION.
CC -1- INDUCTION: BY RETINOIC ACID.
CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY.
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sb.ch/announce/> or send an email to license@isb-sb.ch).
CC -----
CC EMBL; X76290; CAA53922.1; -
CC PDB; 1VHH; 29-JAN-96.
CC MEROPS; C46.002; -
CC MGD; MGI:98297; Shh.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR001767; Hedgehog_hnt.
DR InterPro: IPR003586; Hedgehog_hntc.
DR InterPro: IPR003587; Hedgehog_hntn.
DR InterPro: IPR002203; Intein.
DR InterPro: IPR001657; SonICHH.
DR Pfam; PF01079; Hnt; 1.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHHOG.
DR PRODOM; PD003042; HH_signal; 1.
DR SMART; SM00305; Hntc; 1.
DR SMART; SM00306; Hntn; 1.
DR PROSITE; PS50817; INTEIN_N_TER; 1.

KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal; Lipoprotein; Palmitate; 3D-structure.
 FT CHAIN 1 24 POTENTIAL.
 FT CHAIN 25 437 SONIC HEDGEHOG PROTEIN.
 FT CHAIN 25 198 SONIC HEDGEHOG PROTEIN N-PRODUCT.
 FT CHAIN 199 437 SONIC HEDGEHOG PROTEIN C-PRODUCT.
 FT SITE 198 199 CLEAVAGE (AUTO-) (BY SIMILARITY).
 FT SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY
 FT SITE 244 244 SIMILARITY).
 FT SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY
 FT ACT_SITE 271 271 SIMILARITY).
 FT BINDING 198 198 ESSENTIAL FOR AUTO-CLEAVAGE (BY
 FT DOMAIN 383 387 SIMILARITY).
 FT LIPID 25 25 CHOLESTEROL (BY SIMILARITY).
 FT LIPID 25 25 POLY-GLY.
 SQ SEQUENCE 437 AA; 47773 MW; DOEB72F08E7860DF CRC64;

Query Match 80.7%; Score 1791; DB 1; Length 437;
 Best Local Similarity 80.5%; Pred. No. 8e-138;
 Matches 354; Conservative 21; Mismatches 43; Indels 22; Gaps 6;

QY 4 MLTLRLVLFICALVSSGLTGGPGGICGRHPPKLTPLAKQFIPNAEKTLAGS 63
 2 LLLLRFLVLASSLVCPGLACGPGGFGKRRHPPKLTPLAKQFIPNAEKTLAGS 61
 QY 64 RYEEKITRNSERFELTPNYNPDIIFKDEENTGADRLMTORCKDLNALAISVMNQPGV 123
 DB 62 RYEEKITRNSERFELTPNYNPDIIFKDEENTGADRLMTORCKDLNALAISVMNQPGV 121
 QY 124 KLRVTEGMDDEGHHSESLAYEGRAVDITTSRDRSKYGLMARLAVENGFDWYVESKAH 183
 DB 122 KLRVTEGMDDEGHHSESLAYEGRAVDITTSRDRSKYGLMARLAVENGFDWYVESKAH 181
 QY 184 IHCVKAKENSAVNAASGCGFPSSATVHLEHGCTKLVKDLSPEDRYLADADRLYSDEL 243
 DB 182 IHCVKAKENSAVNAASGCGFPSSATVHLEHGCTKLVKDLSPEDRYLADADRLYSDEL 241
 QY 244 FLDMDSRRKLFYVETKOPARLLTFAHLTFVAPOHNOSEATGTSFG-QALFASVKNP 302
 DB 242 FLDMDSRRKLFYVETKOPARLLTFAHLTFVAPOHNOSEATGTSFG-QALFASVKNP 297
 QY 303 GQRYVYVAGE--GGQQLLPASVSHVSLREASGAYAPLTAAGTIIINRYLASCYAVIEEHS 360
 DB 298 GQRYVYVAGE--GGQQLLPASVSHVSLREASGAYAPLTAAGTIIINRYLASCYAVIEEHS 357
 QY 361 WAHAPAFPRFRLAAGLIALACP-----DGAIPRA-----NTTIGIMYSLLYRI 405
 DB 358 WAHAPAFPRFRLAAGLIALACP-----DGAIPRA-----NTTIGIMYSLLYRI 417
 QY 406 GSWVLDGDALHPLGMVAPAS 425
 DB 418 GSWVLDGDALHPLGMVAPAS 437

RESULT 4
 ID SHH_RAT STANDARD; PRT; 437 AA.
 AC 063673;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sonic hedgehog protein precursor (SHH).
 GN SHH OR VHH-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eulalia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SURAIN-Sprague-Dawley; TISSUE=Embryonic floor plate;
 RA MEDLINE=94170375; PubMed=8124714;
 RA Koellink H., Augsburg A., Heemskerk J., Korzh V., Norlin S.,

RA Ruiz i Altaba A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,
 RA Dodd J.,
 RA "Floor plate and motor neuron induction by vhh-1, a vertebrate homolog
 of hedgehog expressed by the notochord.";
 RL Cell 76:761-775(1994).
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
 CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
 CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE NODE, NOTOCHORD, FLOOR PLATE,
 CC AND POSTERIOR LIMB BUD MESENCHYME.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).
 CC
 CC EMBL: L27340; AAA20999.1;
 CC HSP: O62226; 1VHH.
 CC MEROPS: C46.002;
 DR InterPro: IPR000320; HH_signal.
 DR InterPro: IPR001767; Hedgehog_hlnt.
 DR InterPro: IPR003586; Hedgehog_hlntc.
 DR InterPro: IPR003587; Hedgehog_hlntc.
 DR InterPro: IPR002203; Intein.
 DR InterPro: IPR001657; SonichH.
 DR Pfam: PF01079; Hlntc.1.
 DR Pfam: PF01085; HH_signal.1.
 DR PRINTS: PR00632; SONICHOG.
 DR PRODOM: PD003042; HH_signal.1.
 DR SMART: SM00305; Hlntc.1.
 DR SMART: SM00306; Hlntn.1.
 DR PROSITE: PS50817; INTEIN_N_TER.1.
 DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal; Lipoprotein; Palmitate.
 FT CHAIN 1 24 POTENTIAL.
 FT CHAIN 25 437 SONIC HEDGEHOG PROTEIN.
 FT CHAIN 25 198 SONIC HEDGEHOG PROTEIN N-PRODUCT.
 FT CHAIN 199 437 SONIC HEDGEHOG PROTEIN C-PRODUCT.
 FT SITE 198 199 CLEAVAGE (AUTO-).
 FT SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY
 FT SITE 244 244 SIMILARITY).
 FT SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY
 FT ACT_SITE 271 271 SIMILARITY).
 FT BINDING 198 198 ESSENTIAL FOR AUTO-CLEAVAGE (BY
 FT BINDING 198 198 CHOLESTEROL (BY SIMILARITY).

```

FT DOMAIN 383 387 POLY-GLY.
FT LIPID 25 25 PALMITATE (BY SIMILARITY).
SQ SEQUENCE 437 AA: 47630 MW: 0DBEC19F0D1662A0 CRC64;

Query Match
Best Local Similarity 80.3%; Score 1780; DB 1; Length 437;
Matches 352; Conservative 22; Mismatches 44; Indels 22; Gaps 6;

QY 4 MLITFLVLFICALVSSGLTCGPGRGIGKRRHKKLPFLAVKQFIPVNAKKTIGASG 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 LLLALCFVALVALSLVYCGIACGPGRGKQHPKRLPLAVKQFIPVNAKKTIGASG 61

QY 64 RYGGKTTIRNSERKELTPNNPDIIFKDEBNTGADRLMTORCKDKLNALAIYVNMQPGV 123
DB 62 RYGGKTTIRNSERKELTPNNPDIIFKDEBNTGADRLMTORCKDKLNALAIYVNMQPGV 121

QY 124 KLVTGMDGDGHHSFSLHYEGRAVDITTSDBDRSKYGMRLAVAGFDWYVESKAR 183
DB 122 KLVTGMDGDGHHSFSLHYEGRAVDITTSDBDRSKYGMRLAVAGFDWYVESKAR 181

QY 184 IHCSVAENSVAAKSGCGPFSATVHLHSGTKLVKDLSPGDRVLAADAGRLYSDFLT 243
DB 182 IHCSVAENSVAAKSGCGPFSATVHLHSGTKLVKDLSPGDRVLAADAGRLYSDFLT 241

QY 244 FLDRMDSRKLEFYVETROPARLLTPAHLTVAPQHNSGATGSGQA-LFASNWK 302
DB 242 FLDRMDSRKLEFYVETROPARLLTPAHLTVAPQHNSGATGSGQA-LFASNWK 297

QY 303 GORVYVGE--GGQOLLPASVHSVSLREASGAVAPLTACGTLINNVLASCAVIEHS 360
DB 298 GORVYVGE--GGQOLLPASVHSVSLREASGAVAPLTACGTLINNVLASCAVIEHS 357

QY 361 MAHAPAPFLAAGLALALCP-----DGAIP-----AATTTGTHWSRLYRI 405
DB 358 MAHAPAPFLAAGLALALALPARTDGGGSGIPAPQSVANARACGAPAGIHWSQLLYHI 417

QY 406 GSNVLDGDLHPRGVAPAS 425
DB 418 GTWLIDSETLHPGLMAVKSS 437

RESULT 5
SHH_CYNPY STANDARD: PRT; 432 AA.
AC Q90385;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sonic hedgehog protein precursor (SHH).
GN SHH.
OS Cynops pyrrhogaster (Japanese common newt).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidae; Salamandridae; Cynops.
OX NCBI_TaxID=8330;

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=96136334; PubMed=8573168;
RA Takahashi T., Takahashi T.C., Inoue K., Ogawa M., Takeshima K.;
RT *Activation of two Cynops genes, fork head and sonic hedgehog, in
   animal cap explants.*
RL Blochm. Biophys. Res. Commun. 218:395-401(1996).

-1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
   PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED
   IN LIMB FORMATION, PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND
   VENTRAL SOMITE DIFFERENTIATION. INDUCES ECTOPIC CEMENT GLAND
   FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH
   FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE
   TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC
   REPRESENTS THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
   SIMILARITY).
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
   CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE

```

```

CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- INDUCTION: ACTIVATED BY ACTIVIN, BASIC FIBROBLAST GROWTH FACTOR
CC (BFGF) AND FORK HEAD.
CC -1- PPM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPHAGOLYSIS ACTIVITY
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D63339; BAA09657.1; -.
DR HSSP: Q62226; 1VHH.
DR MEROPS: C46.002; -.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR001767; Hedgehog_hntc.
DR InterPro: IPR003586; Hedgehog_hntc.
DR InterPro: IPR003587; Hedgehog_hntc.
DR InterPro: IPR002203; Intein.
DR InterPro: IPR001657; SonicHH.
DR Pfam: PF01079; Hntc; 1.
DR Pfam: PF01085; HH_signal; 1.
DR PRINTS: PD00632; SONICHHOG.
DR PRODOM: PD003042; HH_signal; 1.
DR SMART: SM00305; Hntc; 1.
DR SMART: SM00306; Hntc; 1.
DR PROSITE: PS50817; INTEIN_N_TER; 1.
DR Developmental protein: Autocatalytic cleavage; Hydrolyase; Protease;
KW signal; Lipoprotein; Palmitate.
FT SIGNAL 1 26
FT CHAIN 27 432
FT CHAIN 27 200
FT CHAIN 201 432
FT SITE 200 201
FT SITE 268 268
FT ACT_SITE 271 271
FT BINDING 200 200
FT LIPID 27 27
SQ SEQUENCE 432 AA: 47847 MW: B455C7E746C8E5A8 CRC64;

Query Match
Best Local Similarity 72.3%; Score 1603.5; DB 1; Length 432;
Matches 316; Conservative 36; Mismatches 70; Indels 13; Gaps 5;

QY 1 MYVMLITRLVLFICALVSSGLTCGPGRGIGKRRHKKLPFLAVKQFIPVNAKKTIG 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 MDEMILRLRYLAFGLICALLVPGLSCGPGRGIGTKRKKLPFLAVKQFIPVNAKKTIG 60

QY 61 ASRGYGGKTTIRNSERKELTPNNPDIIFKDEBNTGADRLMTORCKDKLNALAIYVNMQ 120
DB 61 ASRGYGGKTTIRNSERKELTPNNPDIIFKDEBNTGADRLMTORCKDKLNALAIYVNMQ 120

QY 121 PGVKLVTEGMDGDGHHSFSLHYEGRAVDITTSDBDRSKYGMRLAVAGFDWYVES 180
DB 121 PGVKLVTEGMDGDGHHSFSLHYEGRAVDITTSDBDRSKYGMRLAVAGFDWYVES 180

QY 181 KAHIHCSVAENSVAAKSGCGPFSATVHLHSGTKLVKDLSPGDRVLAADAGRLYSD 240
DB 181 KAHIHCSVAENSVAAKSGCGPFSATVHLHSGTKLVKDLSPGDRVLAADAGRLYSD 240

```

```

OY 241 FLTLDRDSSRLFYIETROPARLLIAHLFVAPQH--NOS-----EATGTSIGQ 293
D 241 FLTLMDKEEYRKVEYVETHSRERVR--LTAHLLFVQAHGNDSCGDFRSVSGAFR 298
OY 294 ALFASNKPGQRYVVLGGGQGLLPASVHSYSLREASGAAPLAQTILLINVLASY 353
D 299 SFEASSVRAGRVLTVDREGGLREAVTERVYL--EATGATAPYATGCTVIDRVLASY 357
OY 354 AVIEHSHMAHAFAPRLAQLALCP---DGAIPYATTTGTGIMHYSRLLYRIGSVWL 410
D 358 AVIEHSHMAHAFAPRLAQLGILISFSPQDSSHSPPRSGEYHWSLILYRIGTVWL 417
OY 411 DGDALHPLGVNAPAS 425
D 418 QEDTIHPLGMAAKSS 432

RESULT 6
SHH_XENLA
ID SHH_XENLA STANDARD; PRT; 444 AA.
AC Q92000; Q91894;
DT 15-JUN-1999 (Rel. 38, Created)
DT 15-JUN-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sonic hedgehog protein precursor (X-SHH) (VHH-1).
GN SHH.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
OC Xenopodidae; Xenopus.
OX NCBI_taxid=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=95357169; PubMed=7630736;
RA Stoclov M.A., Shi Y.-B.;
RT "Xenopus sonic hedgehog as a potential morphogen during embryogenesis
RT and thyroid hormone-dependent metamorphosis.";
RL Nucleic Acids Res. 23:2555-2562(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=95401852; PubMed=7671800;
RA Ekker S.C., McGrew L.L., Lai C.-J., Lee J.-J., von Kessler D.P.,
RA Moon R.T., Beachy P.A.;
RT "Distinct expression and shared activities of members of the hedgehog
RT gene family of Xenopus laevis.";
RL Development 121:2337-2347(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Notochord;
RX MEDLINE=96028338; PubMed=7551564;
RA Rutz I Altaha A., Jessell T.M., Roelink H.;
RT "Restrictions to floor plate induction by hedgehog and winged-helix
RT genes in the neural tube of frog embryos.";
RL Mol. Cell. Neurosci. 6:106-121(1995).
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
CC PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED
CC IN LIMB FORMATION, PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND
CC VENTRAL SOMITE DIFFERENTIATION. INDUCES ECTOPIC CEMENT GLAND
CC FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH
CC TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC
CC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN NOTOCHORD AND NEURAL
CC FLOOR PLATE DURING EMBRYOGENESIS. IN TADPOLE, HIGH EXPRESSION IS
CC OBSERVED IN PANCREAS/STOMACH, MODERATE EXPRESSION IN TAIL, AND LOW

```

```

CC CC EXPRESSION IN INTESTINE, BRAIN, AND HIND LIMB.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT THE NEURULA (STAGES 16-17).
CC FIRST PEAK OF EXPRESSION AROUND TADPOLE HATCHING (STAGES 33-40).
CC HIGH EXPRESSION OBSERVED IN INTESTINE AT THE CLIMAX OF
CC MORPHOGENESIS (STAGES 60-62) WHEN INTESTINE EPITHELIAL UNDERGOES
CC MORPHOGENESIS.
CC -1- INDUCTION: BY THYROID HORMONE.
CC -1- FUNCTION: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L39213; AAC42227.1; -.
DR EMBL; U26314; AAA85162.1; -.
DR EMBL; L35248; AAA49981.1; -.
DR HSSP; Q62226; IVHH.
DR MEROPS; C46_002; -.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001767; Hedgehog_hntc.
DR InterPro; IPR003586; Hedgehog_hntc.
DR InterPro; IPR003587; Hedgehog_hntcn.
DR InterPro; IPR002203; Intein.
DR InterPro; IPR001657; SonicHH.
DR Pfam; PF01079; Hntc_1.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHHOG.
DR PRODOM; PD003042; HH_signal; 1.
DR SMART; SM00305; Hntc_1.
DR SMART; SM00306; Hntcn_1.
DR PROSITE; PS50817; INTEIN_N_TER; 1.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW Signal; Lipoprotein; Palmitate; Repeat.
FT SIGNAL 1 24
FT CHAIN 25 444
FT CHAIN 25 198
FT CHAIN 199 444
FT SITE 198 199
FT SITE 266 266
FT ACT_SITE 269 269
FT DOMAIN 386 409
FT REPEAT 386 393
FT REPEAT 394 401
FT REPEAT 403 409
FT BINDING 198 198
FT LIPID 25 25
FT CONFLICT 5 9
FT CONFLICT 302 319
FT CONFLICT 432 432
FT CONFLICT 494 494
FT SEQUENCE 444 AA; 49453 MW; 73B4E4932FA2EF2 CRC64;
Query Match Score 1533; DB 1; Length 444;
Best local similarity 66.7%; Pred. No. 7.1e-117;
Matches 301; Conservative 41; Mismatches 73; Indels 36; Gaps 5;
OY 4 MULTIR-ILVGFQICALLVSSGLTCGPGRGIGKRRHPRKPLPLAYKOPIPNVAERTIGAS 62

```



```

DB 1 MVAATQSLLLSFICLYLPACGPGRGICKRRHPKLTPLAYKOFIPNAEKTIGAS 60
OY 63 GYREGKITNSERKELFNYPNDITFKDEENTGADRLMTOCKDKLNLAISVMQWFG 122
DB 61 GREGKITNSOCEKLTNPNDFIMFKDEESTGADRLMTOCKDKLNLAISVMQWFG 120
OY 123 VLKRVTEGWEDGHNSESLHREGRAVDITTSRDRSKYGMALARLAVEAGFDVVEYSKA 182
DB 121 VLKRVTEGWEDGHNSESLHREGRAVDITTSRDRSKYGMALARLAVEAGFDVVEYSKA 180
OY 183 HHCSTKAKNSVAAGSGGCFPGSATVHLHGGTKLVKDISPDGVLAADADGRLYSDFL 242
DB 181 HHCSTKAKNSVAAGSGGCFPGSATVHLHGGTKLVKDISPDGVLAADADGRLYSDFL 240
OY 243 TFLDRDSSRKLFYVETEPORARLLTAAHLLEFVAPOHNSOPEATGSTGOALFASNVNP 302
DB 241 MFIIDRDKVKKLFYVETESQKRLR--LTAHLLLEFVA-----QTKVNGTRFSKVSFASNIOP 294
OY 303 GORVVYLAGGGGOQLFPASVHSVSLREASGAYAPLTAOGTILINRYLASCYAVIEESHMA 362
DB 295 GDLIVADPKTMTLKAVKVEKVDL--BEDTGAVAPLTAHGVVVDQVLASCYAVIEESHMA 353
OY 363 HMAFAPRLAAGLALACDGAIFPNAAT-----TTGG 394
DB 354 HMAFAPRLAAGLALACDGAIFPNAAT-----TTGG 394
OY 395 IHWSRLYRIGSWVLDDGALHPLGMVAPAS 425
DB 414 IHWSRLYRIGSWVLDDGALHPLGMVAPAS 444

```

RESULT 7
SHH_BRARE
ID SHH_BRARE STANARD: PRT: 418 AA.
AC 092008: 013170: 013171;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-OCT-2001 (Rel. 40, Last annotation update)
DE Sonic hedgehog protein precursor (SHH) (VHH-1).
GN SHH OR VHH1.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Embryo;
RA MEDLINE-94170375; PubMed-8124714;
RA Roelink H., Augsburg A., Heemskerk J., Korzh V., Norlin S.,
RA Dudd J., Altiba A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,
RA Rold J.;
RT "Floor plate and motor neuron induction by vhh-1, a vertebrate homolog
RT of hedgehog expressed by the notochord";
RT Cell 76:761-775(1994).
RN [2]
RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
RC MEDLINE-96014264; PubMed-7583153;
RA Ekker S.C., Ungar A.R., Greenstein P., von Kessler D.P., Porter J.A.,
RA Moon R.T., Beachy P.A.;
RT "Patterning activities of vertebrate hedgehog proteins in the
RT developing eye and brain";
RT Curr. Biol. 5:944-955(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE-96083328; PubMed-7579523;
RA Fietz M.J., Concoctet J.-P., Barbosa R., Johnson R., Krauss S.,
RA McMahon A.P., Tablin C., Ingham P.W.;
RT "The hedgehog gene family in Drosophila and vertebrate development";
RT development Suppl. 43-51(1994).
RN [4]
RP SEQUENCE FROM N.A.

```

RX MEDLINE-99225458; PubMed-10207136;
RA Muller F., Chang B., Albert S., Fischer N., Tora L., Strahle U.;
RT "Intrinsic enhancers control expression of zebrafish sonic hedgehog in
RT floor plate and notochord";
RT development 126:2103-2116(1999).
RN [5]
RP SEQUENCE OF 30-92 AND 113-170 FROM N.A.
RC TISSUE-Muscle;
RA MEDLINE-97075114; PubMed-8917540;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
RT closely related to the zebrafish";
RT Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
RN [6]
RP PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
RC NOCCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING
DB OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS
OY FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC)
DB RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO
OY ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,
DB PTC REPRESENTS THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
OY SIMILARITY).
DB
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL
CC TUBE AND BRAIN. ALSO FOUND IN THE NOCCHORD AND IN DEVELOPING FIN
CC BUD. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN DOMAINS THAT
CC INCLUDE A DISCRETE REGION IN THE FLOOR OF THE DIENCEPHALON.
CC
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE IN THE INNER CELL LAYER OF
CC THE EMBRYONIC SHIELD DURING GASTRULATION. BY 9.5 HRS OF
CC DEVELOPMENT, EXPRESSED IN A CONTINUOUS BAND THAT EXTENDS FROM THE
CC TAIL TO THE HEAD, THE ANTERIOR BOUNDARY OF EXPRESSION BEING
CC POSITIONED IN THE CENTER OF THE ANIMAL POLE ANTERIOR TO THE
CC PRESUMPTIVE MIDBRAIN.
CC
CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DB EMBL: L27585; AAA20998.1; -
DB EMBL: U30711; AAC59742.1; -
DB EMBL: Z35669; CAAG4738.1; -
DB EMBL: AF124382; AAD47913.1; -
DB EMBL: U51351; AAB38575.1; -
DB EMBL: U51370; AAB38593.1; -
DB HSSP: Q62226; IVHH.
DB MEROPS: C46.002; -
DB ZFIN: ZDB-GENE-980526-166; shh.
DB InterPro: IPR000320; HH_signal.
DB InterPro: IPR001767; Hedgehog_hntc.
DB InterPro: IPR003586; Hedgehog_hntc.
DB InterPro: IPR003587; Hedgehog_hntc.
DB InterPro: IPR002203; Intein.
DB InterPro: IPR001657; SonichH.
DB Pfam: PF01079; Hnt; 1.
DB Pfam: PF01085; HH_signal; 1.
DB PRINTS: PR00632; SONICHOG.

```


DR	ProDom:PP0030342; HH_signal: 1.
DR	SMART: SMO0305; HincC: 1.
DR	SMART: SMO0306; HincN: 1.
DR	PROSITE: PS50817; INTER: 1.
KW	Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW	Signal; Lipoprotein; Palmitate.
FT	SIGNAL 1 23
FT	CHAIN 24 418
FT	CHAIN 24 197
FT	CHAIN 198 418
FT	CHAIN 197 198
FT	SITE 243 243
FT	SITE 267 267
FT	SITE 267 267
FT	ACT_SITE 270 270
FT	BINDING 197 197
FT	LIPID 24 24
SO	SEQUENCE 418 AA; 46402 MW; CF00DAFFEDF25795 CRC64;
Query March	
Best Local Similarity 67.4%; Score 1494; DB 1; Length 418;	
Matches 290; Conservative 50; Mismatches 76; Indels 8; Gaps	
QY	4 MLLRLILVGFICALVYSSGLTSGPGRGIGKRRIKPKLRPLAYAKQFIPVNAEKTIGASG 63
Db	1 MRLRLVLLVLSLTLSLVVSGLAGCPRGVGRRIKPKLRPLAYAKQFIPVNAEKTIGASG 60
QY	64 RYEGKITNSRERFELPTNYNPDIIFDEENTGADRLTQKCKLALISVYNOMPGV 123
Db	61 RYEGKITNSRERFELPTNYNPDIIFDEENTGADRLTQKCKLALISVYNMHPGV 120
QY	124 KLRVTEGMDGHNSEESLHAYEGRAVDITTSDRQRSGKMLARLAVAGEFDMVYESKAH 183
Db	121 KLRVTEGMDGHNFEESLHAYEGRAVDITTSDRKSKYGLSRILAVAGFDMVYESKAH 180
QY	184 IHCSYKAENSVAAKSGGCGPESATVNLHNGSTKLVKQLSPEDRYLADADGRLLYSDFLT 243
Db	181 IHCSYKAENSVAAKSGGCGPESALVLSLDGCKAVKRLNPDKVLADSDAGNVLFSDFIM 240
QY	244 FLDRDSSRKLFYUIETPQPARLLTAAHLLFAAPQHNSEATGSGALFASNYKPG 303
Db	241 FTDRDSTTRKRFYUIETQPEYKELTAAHLLFVLNDSTEDLHMT---AAVASVYAG 296
QY	304 QRVYVLGEGGQILLPASVHVSLSREKASGAYARLTAOSTLIDNVLASCAYVIEHSHMAH 363
Db	297 QKVMVVDSSG--LKSIVLYQRIYTEQDGRSPARYTANGTIVYDRILASCAYVIEDGLAH 354
QY	364 WAFAPRLAQGLLALCPDGLPIPAALT--TGTGNTSKLLRIKGSVNYLDGDLNHPGMY 421
Db	355 LAAPAPRLVYVVSFLFQPNSSRSMATLQDEGVHWSRLLYOMGTWLDLSNMLNHPGMS 414
QY	422 APAS 425
Db	415 VNSS 418
RESULT 8	
TMNH_BRAE	STANDARD; PRT; 416 AA.
ID	TMNH_BRAE
AC	090419;
DT	15-JUL-1999 (Rel. 38, Created)
DT	15-JUL-1999 (Rel. 38, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	199y-winkle hedegon protein precursor (TMNH).
GN	TMNH.
OS	Brechydano gelio (zebrafish) (Dano gelio).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC	Cyprinidae; Dario.
NC	NCBI_TaxID=7935;
LN	[1]

RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
RC TISSUE-Embryo;
RX MEDLINE-96014264; PubMed-7583153;
RA Ekker J.C., Ungar A.R., Greenstein P., von Kessler D.P.,
RT Porter J.A., Moon R.T., Beachy P.A.;
RT "patterned activities of vertebrate hedgehog proteins in the
developing eye and brain."
RL Curr. Biol. 5:944-955(1995).
CC -I- FUNCTION: INTRACELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
PATTERNING EVENTS DURING DEVELOPMENT. INVOLVED IN DORSO-VENTRAL
PATTERNING OF THE BRAIN AND IN EARLY PATTERNING OF THE DEVELOPING
EYES.
CC -I- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -I- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL
TUBE AND BRAIN. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN
DOMAINS THAT INCLUDE A DISCRETE REGION IN THE FLOOR OF THE
DIENCEPHALON. NOT DETECTED IN THE NOTOCHORD OR DEVELOPING FIN BUD.
CC -I- PM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING. WHEREAS THE
C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

CC EMBL; U30710; AAC59741.1; -.
DR HSSP; O62226; 1VHH.
DR MEROPS; C46.001; -.
DR ZFIN; ZDB-GENE-980526-41; twbh.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001167; Hedgehog_hntc.
DR InterPro; IPR003586; Hedgehog_hntc.
DR InterPro; IPR003587; Hedgehog_hntn.
DR InterPro; IPR001657; SonicHH.
DR Pfam; PF01079; Hntc_1.
DR Pfam; PF01085; HH_signal_1.
DR PRINTS; PR00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.
DR SMART; SMO0305; Hntc_1.
DR SMART; SMO0306; Hntn; 1.
KW Developmental protein; Autocatalytic cleavage; Hydrolyase; Protease;
KM Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 26
FT CHAIN 27 416
FT CHAIN 27 200
FT CHAIN 201 416
FT SITE 200 201
FT SITE 270 270
FT ACT_SITE 273 273
FT BINDING 200 200
FT LIPID 27 457
FT SEQUENCE 416 AA; 46576 MW; 61EC2218309CEE5 CRC64;

Query Match 65.7%; Score 1458; DB 1; Length 416;
Best Local Similarity 68.8%; Pred. No. 7.9e-111;
Matches 287; Conservative 36; Mismatches 86; Indels 8; Gaps 4;

QY 5 LRLTLLVGFICALLVSSGLTGPGRGIGKRRHPRKLPPLAYKQFIIPNVAEKTGASGR 64
 DB 5 LHLKOPALLFCISLLTLPCLACGPGRGYGRHPRKLPPLAYKQFIIPNVAEKTGASGR 64
 QY 65 YEKTRNSRREKELTPNPNPDIIKDEENTGADRLMTQRCCKLALAIIVMNMOPGVK 124
 DB 65 YEKTRNSRREKELTPNPNPDIIKDEENTGADRLMTQRCCKLALAIIVMNMOPGVK 124
 QY 125 LRTYEGMEDGHSESLHYEGRAVDITTSDRDSKYGMRLARLAVAGEFDMVYESKAHT 184
 DB 125 LRTYEGMEDGHSESLHYEGRAVDITTSDRDSKYGMRLARLAVAGEFDMVYESKAHT 184
 QY 185 HCSVKAENSVAAKSGCGFPGSATVHLHGGTKLVKDLSPEDRYLAADADGRLLYSDFLTF 244
 DB 185 HCSVKAENSVAAKSGCGFPGSATVHLHGGTKLVKDLSPEDRYLAADADGRLLYSDFLTF 244
 QY 245 LDMDSRKLFFVYIETROPARLLTAAHLFVAPOHNOSEATGSGOALFASNVKPGQ 304
 DB 245 LDMDSRKLFFVYIETROPARLLTAAHLFVAPOHNOSEATGSGOALFASNVKPGQ 304
 QY 305 RYVVLGEGGQQLLPASVSHVSLREASGAVAPLTAOCTILINVLASCVAVIEHSHMAHW 364
 DB 299 TYLVWMDTCSLSKSVYKRI-YTEHSGSPAPYANGTIIVDQVLASCVAVIENHMAHW 357
 QY 365 AFAPRLAOGGLAALCPDGAIPPTAATTGTHWYSLRLYIGSWLDGALHPGV 421
 DB 358 AFAPRLAOGGLAALCPDGAIPPTAATTGTHWYSLRLYIGSWLDGALHPGV 413

RESULT 9
 ID INH_CHICK STANDARD: PRT; 408 AA.
 AC 098938;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Indian hedgehog protein precursor (INH).
 GN INH.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI:TaxID=9031;
 RN [1]
 RP MEDLINE=96325423; PubMed=8662546;
 RA Vortkamp A., Lee K., Lanske B., Segre G.V., Kronenberg H.M.,
 RA Tabin C.J.;
 RT Regulation of rate of cartilage differentiation by Indian hedgehog
 RT and PTH-related protein.
 RL Science 273:613-622(1996).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)
 CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO
 CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN
 CC ENDOCHONDRAL OSSIFICATION: MAY REGULATE THE BALANCE BETWEEN GROWTH
 CC AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION
 CC OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN DEVELOPING MIDGUT, LUNG AND
 CC CARTILAGE OF DEVELOPING LONG BONES IN THE LIMB.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOBIOLOGOUS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U58511; AAC60010.1; -
 CC HSP: Q62226; 1VHK.
 CC DR MEROPS: C46.003; -
 CC DR InterPro: IPR000320; HH_signal.
 CC DR InterPro: IPR001767; Hedgehog_hlnc.
 CC DR InterPro: IPR003586; Hedgehog_hlnc.
 CC DR InterPro: IPR003587; Hedgehog_hlnc.
 CC DR InterPro: IPR001657; SonicHH.
 CC DR Pfam: PF01079; Hlnc; 1.
 CC DR Pfam: PF01085; HH_signal; 1.
 CC DR PRINTS: PR00632; SONICHHOG.
 CC DR ProDom: PD003042; HH_signal; 1.
 CC DR SMART: SM00305; Hlnc; 1.
 CC DR SMART: SM00306; Hlnc; 1.
 CC DR SMART: SM00306; Hlnc; 1.
 CC KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 CC Signal; Lipoprotein; Palmitate.
 CC FT SIGNAL 1 23
 CC FT CHAIN 24 408
 CC FT CHAIN 24 198
 CC FT CHAIN 199 408
 CC FT SITE 198 408
 CC FT SITE 244 244
 CC FT SITE 244 244
 CC FT SITE 268 268
 CC FT ACT_SITE 271 271
 CC FT BINDING 198 198
 CC FT LIPID 24 24
 CC SQ SEQUENCE 408 AA; 44829 MW; BA397AE2A9357A24 CRC64;

Query Match 58.4%; Score: 1295.5; DB 1; Length 408;
 Best Local Similarity 64.3%; Pred. No. 1.2e-97;
 Matches 266; Conservative 39; Mismatches 90; Indels 19; Gaps 7;

QY 9 RLLVGFICALLVSSGLTGPGRGIGKRRHPRKLPPLAYKQFIIPNVAEKTGASGR 66
 DB 5 RLLVGFICALLVSSGLTGPGRGIGKRRHPRKLPPLAYKQFIIPNVAEKTGASGR 66
 QY 67 GKTTRNSRREKELTPNPNPDIIKDEENTGADRLMTQRCCKLALAIIVMNMOPGVK 126
 DB 65 GKTTRNSRREKELTPNPNPDIIKDEENTGADRLMTQRCCKLALAIIVMNMOPGVK 124
 QY 127 VTEGMDGHSESLHYEGRAVDITTSDRDSKYGMRLARLAVAGEFDMVYESKAHT 186
 DB 125 VTEGMDGHSESLHYEGRAVDITTSDRDSKYGMRLARLAVAGEFDMVYESKAHT 184
 QY 187 SYKAENSVAAKSGCGFPGSATVHLHGGTKLVKDLSPEDRYLAADADGRLLYSDFLTF 246
 DB 185 SYKAENSVAAKSGCGFPGSATVHLHGGTKLVKDLSPEDRYLAADADGRLLYSDFLTF 244
 QY 247 RMDSRKLFFVYIETROPARLLTAAHLFVAPOHNOSEATGSGOALFASNVKPGQ 306
 DB 245 KEPRALTAFTVETROPARLLTAAHLFVAPOHNOSEATGSGOALFASNVKPGQ 300
 QY 307 YVVLGEGGQQLLPASVSHVSLREASGAVAPLTAOCTILINVLASCVAVIEHSHMAHW 365
 DB 301 YVVLGEGGQQLLPASVSHVSLREASGAVAPLTAOCTILINVLASCVAVIEHSHMAHW 357
 QY 366 AFAPRLAOGGLAALCPDGAIPPTAATTGTHWYSLRLYIGSWLDGALHPGV 421
 DB 358 AFAPRLAOGGLAALCPDGAIPPTAATTGTHWYSLRLYIGSWLDGALHPGV 402

RESULT 10
 IHH_XENLA STANDARD; PRT; 409 AA.
 AC 091612;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Indian hedgehog protein precursor (IHH) (Banded hedgehog protein) (X-BHH).
 GN IHH OR BHH.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodidae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=95401852; PubMed=7671800;
 RA Ekker S.C., McGrew L.L., Lai C.-J., Lee J.J., von Kessler D.P., Moon R.T., Beachy P.A.;
 RT "Distinct expression and shared activities of members of the hedgehog gene family of Xenopus laevis";
 RL Development 121:2337-2347(1995).
 CC -1 FUNCTION: SIGNAL INVOLVED IN THE EARLY INDUCTION AND PATTERNING OF ANTERODORSAL ECTODERM, NERVOUS SYSTEM AND SOMITES. INDUCES ECTOPIC CEMENT GLAND FORMATION IN EMBRYOS.
 CC -1 SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. HEDGEHOG PROTEIN IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1 DEVELOPMENTAL STAGE: DETECTABLE WITHIN THE EARLY GASTRULA. AT STAGE 14 NEURULA, HIGH EXPRESSION IN NEUROCTODERM, EXPRESSED THROUGHOUT THE NEURAL PLATE AND SUBSEQUENTLY IN BOTH THE NERVOUS SYSTEM AND IN THE DERMATOME OF SOMITES.
 CC -1 INDUCTION: BY ACTIVIN.
 CC -1 PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1 SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to license@sdb.ch).
 CC -----
 DR EMBL: U26404; AAA85165.1; .
 DR HSSP: 062226; 1VHH.
 DR MEROPS: C46.UPW.
 DR InterPro: IPR000320; HH_signal.
 DR InterPro: IPR001767; Hedgehog_hlnt.
 DR InterPro: IPR003586; Hedgehog_hlntc.
 DR InterPro: IPR003587; Hedgehog_hlntn.
 DR InterPro: IPR001657; SonichH.
 DR Pfam: PF01079; Hint; 1.
 DR Pfam: PF01083; HH_signal; 1.
 DR PRINTS: PR00632; SONICHOG.
 DR PRODOM: PD003042; HH_signal; 1.
 DR SMART: SM00305; Hlntc; 1.
 DR SMART: SM00306; Hlntn; 1.
 KM Developmental protein; Autocatalytic cleavage; Hydrolase; Protease; Signal; Lipoprotein; Palmitate.

FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 409 INDIAN HEDGEHOG PROTEIN, N-PRODUCT.
 FT CHAIN 24 197 INDIAN HEDGEHOG PROTEIN C-PRODUCT.
 FT CHAIN 198 409 INDIAN HEDGEHOG PROTEIN C-PRODUCT.
 FT SITE 197 198 CLEAVAGE (AUTO-) (BY SIMILARITY).
 FT SITE 267 267 INVOLVED IN AUTO-CLEAVAGE (BY SIMILARITY).
 FT ACT_SITE 270 270 ESSENTIAL FOR AUTO-CLEAVAGE (BY SIMILARITY).
 FT BINDING 197 197 CHOLESTEROL (BY SIMILARITY).
 FT DOMAIN 32 35 POLY-ARG. (BY SIMILARITY).
 FT LIPID 24 24 PALMITATE (BY SIMILARITY).
 SQ SEQUENCE 409 AA: 45591 MW: 66B265367FB96827 CRC64.
 Query Match 58.0%; Score 1286; DB 1; Length 409;
 Best Local Similarity 61.2%; Pred. No. 7e-97;
 Matches 256; Conservative 54; Mismatches 90; Indels 18; Gaps 6;
 QY 7 LTRILLVGFICALVSSGILT-CGPRGIGRRHPPKRLPLAYKQFIPIVNAERTLAGSGRY 65
 DB 3 LPRVYLLCAALLLSGAVRGCGPRVGRRRRPTLSLSPKQSPNPEKTLGASGRY 62
 QY 66 ECKTRNSRFEKELPPNYPDIIFKDENTGADRLTQCKRKLNALAISVNMQPGVRL 125
 DB 63 EKRISNSRFEKELPPNYPDIIFKDEETGADRLTQCKRKLNSLAISVNMQPGVRL 122
 QY 126 RTEGDEDEHSESESLHYEGRAVDITSDRPSKGMRLAYEGPMPVYESAHTH 185
 DB 123 RTEGDEDEHSESESLHYEGRAVDITSDRPSKGMRLAYEGPMPVYESAHTH 182
 QY 186 CSVKENSVAANSKGGCFPGSATVLEHGSKLKVDSPLDRLVADADGRLLYDFLPL 245
 DB 183 CSVKESHAANKTGGCFPEALATLESSEKIPVQSPLAVLADNSRPTYSDFLSL 242
 QY 246 DMDSSRLFYIETROPARLLTPAHLTPAPQNSEALGTSIGALLPASNKPGGR 305
 DB 243 DSDPKKEHFOYIKTQDPHRLFLTPAHLLFVSDNT---SPASFOAVFASVAPGO- 297
 QY 306 VYVLGGGQQLLPASVSHSLREBSGAAVPLAOTTLINRVLSCVAVIEHSAHNA 365
 DB 298 -YILVSNVGLLPKAVRSYNTQTN-YGAYAPLTOHGTLLVDDVVVSCFLVQORLQIV 355
 QY 366 FAFRLAQ--GLAALCPDGAIPATTTTGIHWYSRLYRIGSVNLDADALPLGMV 421
 DB 356 YMLPLRVLYMLGIAG-----TQPSQMGIMHWSKALYHIGRLIHQNEHPGLIV 405
 RESULT 11
 IHH_MOUSE STANDARD; PRT; 411 AA.
 AC P97812; 061724;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Indian hedgehog protein precursor (IHH) (HhG-2).
 GN IHH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
 RC STRAIN=CD-1; TISSUE=Kidney;
 RX MEDLINE=97236802; PubMed=9079674;
 RA Valentini R.P., Brocksamer W.T., Park J., Yang T., Briggs J., Dressler G., Holzman L.B.;
 RT "Post-translational processing and renal expression of mouse Indian hedgehog";
 RL J. Biol. Chem. 272:8466-8473(1997).
 RN [2]
 RP SEQUENCE OF 76-411 FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=94094334; PubMed=7916661;

RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,
 RA McMahon J.A., McMahon A.P.;
 RT "Sonic hedgehog, a member of a family of putative signaling
 RT molecules, is implicated in the regulation of CNS polarity.";
 RL Cell 75:1417-1430(1993).
 RN [3]
 RP REVISIONS.
 RC STRAIN=C57BL/6J;
 RA St Jacques B.;
 RL Submitted (APR-1997) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 124-172 FROM N.A.
 RX MEDLINE=95236997; Pubmed=7720571;
 RA Chang D.T., Lopez A., von Kessler D.P., Chang C., Simandl B.K.,
 RA Zhao R., Seidlin M.F., Fallon J.F., Beachy P.A.;
 RT "Products, genetic linkage and limb patterning activity of a murine
 RT hedgehog gene";
 RL Development 120:3339-3353(1994).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)
 CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO
 CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN
 CC ENDOCHONDRAL OSSIFICATION: MAY REGULATE THE BALANCE BETWEEN GROWTH
 CC AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION
 CC OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, DETECTED IN THE DEVELOPING GUT,
 CC AND UROGENITAL SINDS. IN THE ADULT KIDNEY, FOUND IN PROXIMAL
 CC CONVOLUTED AND PROXIMAL STRAIGHT TUBULE.
 CC -1- DEVELOPMENTAL STAGE: DETECTED AT 10 DAYS POST COITUM (DPC) IN
 CC IN THE DEVELOPING UROGENITAL SINDS. EXPRESSION INCREASES WITH
 CC GESTONAL AGE IN KIDNEY AND DUODENUM, BECOMING MAXIMAL IN
 CC ADULTHOOD.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U85610; AAB49692.1; ALT_INIT.
 DR EMBL; X76291; CA53923.1; -.
 DR HSSP; Q62226; LVHH.
 DR MEROPS; C46.003; -.
 DR MGD; MGI:96533; Inh.
 DR InterPro; IPR000320; HH_signal.
 DR InterPro; IPR001767; Hedgehog_hint.
 DR InterPro; IPR003586; Hedgehog_hintc.
 DR InterPro; IPR003587; Hedgehog_hintn.
 DR InterPro; IPR002203; Intein.
 DR Pfam; PF01079; Hinc; 1.
 DR Pfam; PF01085; HH_signal; 1.
 DR ProDom; PD003042; HH_signal; 1.
 DR SMART; SM00305; Hinc; 1.
 DR SMART; SM00306; Hinc; 1.
 DR PROSITE; PS50817; INTEIN_N_TER; 1.

KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KM Signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 27
 FT CHAIN 28 411
 FT CHAIN 28 202
 FT CHAIN 203 411
 FT SITE 202 203
 FT SITE 202 203
 FT SITE 248 248
 FT SITE 272 272
 FT ACT_SITE 275 275
 FT BINDING 202 202
 FT LIPID 28 28
 FT CONFLICT 383 383
 FT SEQUENCE 411 AA; 45485 MW; 08BEFADB507C00DB CRC64;
 Query Match 57.68; Score 1277.5; DB 1; Length 411;
 Best Local Similarity 62.48; Pred. No. 3,4e-96;
 Matches 257; Conservative 42; Mismatches 88; Indels 25; Gaps 6;
 QY 19 LVSSGLTCGGRCIG-KRRHPRKLTPLAYKQFIPNVAEKTLAGSVECKITNSERFK 77
 DB 20 LTVPAKCGCGRVGSSRRPRRLVPLAKQSPNPEKTLGSGREGKIAHSSERFK 79
 QY 78 ELTPNYPNDIFKDEWNTGADRLTQRCOKDLNLAISVNMQGVKLRVTEGDEGHN 137
 DB 80 ELTPNYPNDIFKDEWNTGADRLTQRCOKDLNLAISVNMQGVKLRVTEGDEGHN 139
 QY 138 SEESLHTEGRAVDITTSDDRSKTYGMLARLAVENGFWYVESFAHHCYKAENVAAK 197
 DB 140 SEESLHTEGRAVDITTSDDRSKTYGMLARLAVENGFWYVESFAHHCYKAENVAAK 199
 QY 198 SGGEFPGSAFTHLHGKTKYKQDLSPPDRVLAADADRLYSDTLTDRDSSKLFY 257
 DB 200 TGGEFPGSAFTHLHGKTKYKQDLSPPDRVLAADADRLYSDTLTDRDSSKLFY 259
 QY 258 IETOPPARLLTAHLFLFVAHQHNSGANGSTGQALFASNVPGORVYVLEGGOQL 317
 DB 260 IETOPPARLLTAHLFLFVAHQHNSGANGSTGQALFASNVPGORVYVLEGGOQL 313
 QY 318 PASVHSLKREASGAVAPLTAOCTILINRYLASCIVAVIEHSVAHNAFAPRLAQGLA 377
 DB 314 PARVAASV-THVALGSYAPLFRHGTLVVEDVVASCFPAVADHHAOLAFLPRL- 366
 QY 378 ALCPDGAIPFAA-----TTTGTHVYSRLYRIGSWVLDGQALHPLGVAPAS 425
 DB 367 -----FPPLAMGSMTPBSGVNTPQWLYRGLLLEESTFPLGSGAGS 411
 RESULT 12
 ID IHH_HUMAN STANDARD; PRT; 411 AA.
 AC Q14623; Q43322;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Indian hedgehog protein precursor (IHH) (HNG-2).
 GN IHH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tate G., Kishimoto K., Mitsuya T.;
 RT "Expression of Sonic hedgehog and its receptor Patched/Smoothed in
 RT human cancer cell lines and embryonic organs";
 RL J. Biochem. Mol. Biol. Biophys. 4:27-34(2000).
 RP [2]
 RP SEQUENCE OF 100-411 FROM N.A.
 RC TISSUE=Fetal lung;

RX MEDLINE-96070431; PubMed-7590746;
 RA Marigo V., Roberts D.J., Lee S.M.K., Tsukurov O., Levy T.,
 RA Gaslier J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seidman C.E.,
 RA Jenkins N.A., Seidman J.G., McMahon A.P., Tabin C.;
 RT "Cloning, expression, and chromosomal location of SHN and INH: two
 RT human homologues of the Drosophila segment polarity gene hedgehog.";
 RN Genomics 28:44-51(1995).
 RN (3)
 RP SEQUENCE OF 124-172 FROM N.A.
 RX MEDLINE-95236997; PubMed-7720571;
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
 RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;
 RT "Products, genetic linkage and limb patterning activity of a murine
 RT hedgehog gene.";
 RL Development 120:3339-3353(1994).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)
 CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO
 CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN
 CC ENDOCHONDRAL OSSIFICATION. MAY REGULATE THE BALANCE BETWEEN GROWTH
 CC AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION
 CC OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP) (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN EMBRYONIC LUNG, AND IN ADULT
 CC KIDNEY AND LIVER.
 CC -1- PM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING. WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AB018076; BAA33523.2; -
 DR EMBL; AB010092; BAA33523.2; JOINED.
 DR EMBL; AB018075; BAA33523.2; JOINED.
 DR EMBL; L38517; AAG62178.1; -
 DR HSSP; Q62226; 1VHH.
 DR MEROPS; C46.003; -
 DR GeneW; HGNC:5956; INH.
 DR MIM; 600726; -
 DR InterPro; IPR000320; HH_signal.
 DR InterPro; IPR001767; Hedgehog_hint.
 DR InterPro; IPR003586; Hedgehog_hint.
 DR InterPro; IPR003587; Hedgehog_hintn.
 DR InterPro; IPR002203; Inteln.
 DR Pfam; PF01079; Hint; 1.
 DR Pfam; PF01085; HH_signal; 1.
 DR SMART; SM00305; Hintc; 1.
 DR SMART; SM00306; Hintn; 1.
 DR PROSITE; PS50817; INTERIN_N_TER; 1.
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 27
 FT CHAIN 28 411
 FT CHAIN 28 202
 FT CHAIN 203 411
 FT SITE 202 203
 FT SITE 248 248
 FT SITE 248 248
 INVOLVED IN CHOLESTEROL TRANSFER (BY
 POTENTIAL.
 INDIAN HEDGEHOG PROTEIN.
 INDIAN HEDGEHOG PROTEIN N-PRODUCT.
 INDIAN HEDGEHOG PROTEIN C-PRODUCT.
 CLEAVAGE (AUTO-) (BY SIMILARITY).
 INVOLVED IN CHOLESTEROL TRANSFER (BY

FT SITE 272 272
 FT ACT_SITE 275 275
 FT BINDING 202 202
 FT LIPID 28 28
 FT CONFLICT 100 100
 FT CONFLICT 246 246
 FT CONFLICT 309 309
 SQ SEQUENCE 411 AA; 45188 MW; 14C904A3C344DEA CRC64;
 Query Match 57.1%; Score 1265.5; DB 1; Length 411;
 Best Local Similarity 62.7%; Pred. No. 3.2e-95;
 Matches 256; Conservative 38; Mismatches 97; Indels 17; Gaps 5;
 OY 19 LLYSGGLTCGPRGIG-KRRHKRLPLAAYKPTRNAVKRTIAGSGRREGKTRNSERK 77
 DB 20 LVVPAMGCGPGRVYVSRPRPKLVPLAYKQFSNVPEKTIGASGRYEGKTAERSERK 79
 OY 78 ELTPVNPDIIRKDEENNGADRLMTORCKDLNALAISVNMOPGVKLVTEGMDXHH 137
 DB 80 ELTPVNPDIIRKDEENNGADRLMTORCKDLNSLAISVNMOPGVKLVTEGMDXHH 139
 OY 138 SEESLHYEGRAVDITSDRSKGYMLARLAVEAGFDWYYSKAKIHCSVAENSVAK 197
 DB 140 SEESLHYEGRAVDITSDRSKGYMLARLAVEAGFDWYYSKAKIHCSVAENSVAK 199
 OY 198 SCGCPGSAATVLEHGRTKLVNDLSRQGVNLADADGRLYSDFLTPDRMDSKRLPV 257
 DB 200 TGCCPPAGQVLESGARVALSAVRGDRVLAMGDSPTFSDVILLDREPHRRARFQY 259
 OY 258 ITTPPRARLTAHLLEFAPHOHSEATGSGOALFASNVKPGQRYVVEGGQQL 317
 DB 260 ITTPPRARLTAHLLEFAPHOHSEATGSGOALFASNVKPGQRYVVEGGQQL 319
 OY 318 PASVSVSLREBSAGAYAPLTAQGTILINRVLASCYAVIEESHMAWAFAPRLAOGLLA 377
 DB 314 PARVAAYS-THVALGAYAPLTKHGTIVVEDVVASCFPAVAADHHLQALRWPLRLFLSHLW 372
 OY 378 ALCPDGAITPATTTTGTGWYSRLXLRIGSWLDDGALPRLGVNAPAS 425
 DB 373 G-----SWTPGEGVMYPOLLYRLGRLLRLESGSFPLDMSGAGS 411
 RESULT 13
 ITH_BRARE
 AC 098862; STANDARD; PRT; 412 AA.
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Indian hedgehog protein precursor (INH) (Echidna hedgehog protein).
 GN INH OR EHH.
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_Taxid=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96310864; PubMed-8684485;
 RA Currie P.D., Ingham P.W.;
 RT "Induction of a specific muscle cell type by a hedgehog-like protein
 RT in zebrafish.";
 RL Nature 382:452-455(1996).
 RN [2]
 RP SEQUENCE OF 113-170 FROM N.A.
 RC TISSUE=Muscle;
 RX MEDLINE-97075114; PubMed-8917540;
 RA Zardoya R., Abouheif E., Meyer A.;
 RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
 RT closely related to the zebrafish.";

```

RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
CC PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
CC NOTOCHORD THAT INDUCES SOMITE PATTERNING AND MUSCLE PIONEER
CC DIFFERENTIATION.
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN THE NOTOCHORD.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT THE MID-GASTRULA STAGE.
CC DISAPPEARS AT THE END OF THE SOMITOGENESIS.
CC -1- PFM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPHAGOLYSIS ACTIVITY
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y08426; CAA69702.1; -.
CC DR HSSP; O62226; 1VHH.
CC DR MEROPS; C46; UPM; -.
CC DR ZFIN: ZDB-GENE-980526-135; enh.
CC DR InterPro: IPR000320; HH_signal.
CC DR InterPro: IPR001767; Hedgehog_hntc.
CC DR InterPro: IPR003586; Hedgehog_hntc.
CC DR InterPro: IPR003587; Hedgehog_hntc.
CC DR InterPro: IPR002203; Intein.
CC DR InterPro: IPR001657; SotichH.
CC DR Pfam; PF01079; Hntc; 1.
CC DR Pfam; PF01085; HH_signal; 1.
CC DR PRINTS; PR00632; SONICHOG.
CC DR PRODOM; PD003042; HH_signal; 1.
CC DR SMART; SM00305; Hntc; 1.
CC DR SMART; SM00306; Hntc; 1.
CC DR PROSITE; PSS0817; INTEN_N_TER; 1.
CC DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
CC KW Signal; Lipoprotein; Palmitate.
CC FT SIGNAL 1 23
CC FT CHAIN 24 412
CC FT CHAIN 24 197
CC FT CHAIN 198 412
CC FT SITE 197 198
CC FT SITE 246 246
CC FT SITE 269 269
CC FT ACT_SITE 272 272
CC FT BINDING 197 197
CC FT LIPID 24 24
CC FT CONFLICT 121 121
CC SO SEQUENCE 412 AA; 45746 MW; 21D43F052A278CA1 CRC64;

Query Match 55.6%; Score 1232.5; DB 1; Length 412;
Best Local Similarity 60.9%; Pred. No. 1.5e-92;
Matches 254; Conservative 51; Mismatches 91; Indels 21; Gaps 10;
OY 11 LVVFICALLVSSGLGSGRGIGRRHKKRPLAYKOFIPNVAETIASGVEKIT 70
DB 9 LITGFIAPSPAD-GCGPGRGIGKRRTPKRLPLAYKOPSPNVAETIASGVEKIT 67

```

```

OY 71 RNSRFKELPNYNDILFDEENTGADRLMTQRCCKDLNALAISVNMOPGVRLRYTEG 130
DB 68 PSSRFKELPNYNDILFDEENTGADRLMTQRCCKDLNALAISVNMOPGVRLRYTEG 127
OY 131 WDEGHHSESLHREGAVDITTSDRORSKYGMALRLAVGEFWMYTESKAHHCYKA 190
DB 128 WDEGHHSESLHREGAVDITTSDRORSKYGMALRLAVGEFWMYTESKAHHCYKA 187
OY 191 ENSVAKSGGCFPSATVHLEHGSTKLVDKSPDRVLA---ADADGRLLYSDFLTDR 247
DB 188 EHSVAATGCGCFPRALVTYKMGSHROIQADGLVASGSGTGLTYEVLTFLDR 247
OY 248 MDSRKLELYETROPAPARLLTAHLF--VAPDHQSEATGTSQALFASNVKPGOR 305
DB 248 RPIQKHFVYIRT-EDGASVSLTAHLHFRKGVGNSRGEKPG-AVRTIFASDAQGQ- 304
OY 306 VYVGEQGQQLPASVSVSLREASGAYAPLTAGTITLNRVLASCIYAVIEHSAHMA 365
DB 305 CLLGKLRKR--SQTHVGRED-QGLYPLTAHGTVVNDVLTSCAAVNRQLAHMA 361
OY 366 FAPRLAGLLALCPGAIPTATTTGTGHWYSRLLYRIGSWYLDGDLHPGMA 422
DB 362 FAPRL--LYSWGPDQVL-----KNGLHWISQVLIGKLLDSELPHPALAE 409

RESULT 14
ID DHH_MOUSE STANDARD; PRT; 396 AA.
AC 061488;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Desert hedgehog protein precursor (Dhh) (HHG-3).
GN DHH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=94094334; Pubmed=7916661;
RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,
RA McMahon J.A., McMahon A.P.;
RT "sonic hedgehog, a member of a family of putative signaling
RT molecules, is implicated in the regulation of CNS polarity.";
RL Cell 75:1417-1430(1993).
RN [2]
RP SEQUENCE OF 120-168 FROM N.A.
RX MEDLINE=95236997; Pubmed=7720571;
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
RA Zhao R., Seidlin M.F., Fallon J.F., Beachy P.A.;
RT "Products, genetic linkage and limb patterning activity of a murine
RT hedgehog gene.";
RL Development 120:3339-3353(1994).
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A
CC SPERMATOCTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES
CC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ADULT TESTES. NOT EXPRESSED IN
CC LIMB BUDS.
CC -1- PFM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPHAGOLYSIS ACTIVITY
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE

```

CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 DR EMBL: X76292; CA53924.1; -
 DR HSSP: Q62226; LVH.
 DR MEROPS: C46.004; -
 DR MGD: MGI:94891; Dh.
 DR InterPro: IPR000320; HH_signal.
 DR InterPro: IPR001767; Hedgehog_hnt.
 DR InterPro: IPR003586; Hedgehog_hntC.
 DR InterPro: IPR003587; Hedgehog_hntN.
 DR InterPro: IPR002203; Intein.
 DR Pfam: PF01079; Hint; 1.
 DR Pfam: PF01085; HH_signal; 1.
 DR PRINTS: PR00632; SONICHOG.
 DR PRODOM: PD003042; HH_signal; 1.
 DR SMART: SM00305; HntC; 1.
 DR SMART: SM00306; HntN; 1.
 DR PROSITE: PSS0817; INTEIN_N_TER; 1.
 DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 22
 FT CHAIN 23 396
 FT CHAIN 23 198
 FT CHAIN 199 396
 FT SITE 198 199
 FT SITE 244 244
 FT SITE 268 268
 FT ACT_SITE 271 271
 FT BINDING 198 198
 FT LIPID 23 23
 SQ SEQUENCE 396 AA; 43542 MW; AFEE051BE950FD8 CRC64;
 Query Match 54.18; Score 1200; DB 1; Length 396;
 Best Local Similarity 60.78; Pred. No. 6.3e-90; Indels 16; Gaps 9;
 Matches 241; Conservative 56; Mismatches 84;
 QY 12 LVGFTIC-ALVSSGLTGCGRG-IGKRRHPRK-LTPLAYKQIPNVAEKTLAGSGRYEGK 68
 Db 7 LPLPCCLALLALASOSGPGRGVGRYRKQLVPLLYKQFVSPMERTIGASGAPGR 66
 QY 69 ITRNSERKELTPYNDPIIKDEBNGADLMQRCOKDNALAIISVMNQPGVKLRVT 128
 Db 67 VTRSSERRDLVPPNYNDIIFKDEBNGADLMQRCOKDNALAIISVMNQPGVKLRVT 126
 QY 129 EGMDEDGHSSESLHYEGRAVDITTSDRSRKYGMLARLAEAGPDWYYSKAHIHCSV 188
 Db 127 EGMDEDGHSSESLHYEGRAVDITTSDRSRKYGMLARLAEAGPDWYYSKAHIHCSV 186
 QY 189 KAENSVAKSGCGPPGASVAVLHNGTKLVNDLSPGDVLAADADGRLVSDPLFLDRM 248
 Db 187 KADNSLAVRAGCGPPGASVAVLHNGTKLVNDLSPGDVLAADADGRLVSDPLFLDRM 246
 QY 249 DSSKRLTYVITRPPARLLTAHLFLVAPQHNOSEATGTSQALFASVVKPGQRYTV 308
 Db 247 LQRRASFAVAETERRPRLTLTPMHLVFAA--RGPAPAPGDF--PVFARRLRAGDS--V 300
 QY 309 LGEAGDGLPPSVHSVLSREASAGVAPLTAGQITLIRVASCYAVEEESHMHWAPAP 368
 Db 301 LAPGGLALQPARVARVA-REZAVGVFAPLTAHGTLLVNDVLAACVAVLESHQWHRRAP 359

QY 369 FRLAAGLALCPDGAITPAATTTGIMHYSRLRYI 405
 Db 360 LRLHLA-LGALLPGAV-----QPTGMHYSRLRYRL 390
 RESULT 15
 DHH_HUMAN
 ID DHH_HUMAN STANDARD; PRT; 396 AA.
 AC 043323; 015794;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Desert hedgehog protein precursor (DHH) (HHG-3).
 GN DHH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tate G., Kishimoto K., Mitsuya T.;
 RT "Expression of Sonic hedgehog and its receptor Patched/Smoothed in
 RT human cancer cell lines and embryonic organs".
 RL J. Biochem. Mol. Biol. Biophys. 4:27-34(2000).
 RN [2]
 RP SEQUENCE OF 85-178 FROM N.A.
 RA Tate G., Kishimoto K., Mitsuya T.;
 RT "Human desert hedgehog".
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP DISEASE.
 RA PubMed:11017805;
 RA Umehara F., Tate G., Itoh K., Yamaguchi N., Douchi T., Mitsuya T.,
 RA Osame M.;
 RT "A novel mutation of desert hedgehog in a patient with 46,XY partial
 RT gonadal dysgenesis accompanied by multifocal neuropathy".
 RL Am. J. Hum. Genet. 67:1302-1305(2000).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL. ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A
 CC SPERMATOCTYTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES
 CC DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- DISEASE: Defects in DHH may be the cause of partial gonadal
 CC dysgenesis (PGD) with multifocal neuropathy. PGD is
 CC characterized by the presence of a testis on one side and a streak
 CC or an absent gonad at the other, persistence of Mullerian duct
 CC structures, and a variable degree of genital ambiguity.
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 DR EMBL: AB010994; BAA24866.1; -
 DR EMBL: AB010581; BAA24866.1; JOINED.

DR EMBL: AB010993; BAA24866.1; JOINED.
DR EMBL: U59748; AAB03398.1; -.
DR HSSP: Q62226; 1VHH.
DR MEROPS: C46.004; -.
DR Genew: HGNC:2865; DHH.
DR MIM: 605423; -.
DR MIM: 607080; -.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR001767; Hedgehog_hntc.
DR InterPro: IPR003586; Hedgehog_hntc.
DR InterPro: IPR003587; Hedgehog_hntc.
DR InterPro: IPR001657; SonicHH.
DR Pfam: PF01079; Hntc; 1.
DR Pfam: PF01085; HH_signal; 1.
DR PRINTS: PR00632; SONICHHOG.
DR ProDom: PD003042; HH_signal; 1.
DR SMART: SM00305; Hntc; 1.
DR SMART: SM00306; Hntc; 1.
DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 396 DESERT HEDGEHOG PROTEIN.
FT CHAIN 23 198 DESERT HEDGEHOG PROTEIN N-PRODUCT.
FT CHAIN 199 396 DESERT HEDGEHOG PROTEIN C-PRODUCT.
FT SITE 198 199 CLEAVAGE (AUTO-) (BY SIMILARITY).
FT SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY
FT SITE 244 244 SIMILARITY).
FT SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY
FT SITE 268 268 SIMILARITY).
FT ACT_SITE 271 271 ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT BINDING 198 198 SIMILARITY).
FT LIPID 23 23 CHOLESTEROL (BY SIMILARITY).
FT CONFLICT 177 177 PALMITATE (BY SIMILARITY).
SQ SEQUENCE 396 AA: 43577 MW: 43482.19723 AD5 CRC64:
E -> G (IN REF. 2).

Query Match 54.0%; Score 1198; DB 1; Length 396;
Best Local Similarity 59.4%; Pred. No. 9, 2e-90;

Matches 244; Conservative 56; Mismatches 91; Indels 20; Gaps 10;

QY 4 MLTLFRLVGFIC--ALLVSSGLTGGPGRG-IGKRHPK--LTPLAYKOFIPNVAKTL 59
DB 1 MALLTLFLP---LCCALALLAPRQSCGPRGPRGRRYARKOLVPLLYKQFVGPVERTL 57
QY 60 GASGRYEGKITRNSERKEKELTPYNPDIIFKDEENTGADRLMTQRCCKLMLAISMNQ 119
DB 58 GASGPAEGHVARGSERFRLVPYNPDIIFKDEENSGADRLMTRECKERYNMLAIVMMN 117
QY 120 WPGVKRLRTGEGDEGHSEESIHYGRAVDITTSRDSKSGMLARLAVEAGFDVYVE 179
DB 118 WPGVRLRTGEGDEGHSEESIHYGRAVDITTSRDSKSGMLARLAVEAGFDVYVE 177
QY 180 SKAHISYKAEKSVNAGSGGPGSATVHLEHGTGKLVKDSPGDRLAADADGRLLYS 239
DB 178 SRNHVIVSVKADNSLAVRAGGCEPFGNATVRLNGERKGLRELKRDVWLADSGVVPPT 237
QY 240 DELTFLDRDSSRKLEFYIETROPARALLTLAHLFVAPQHNQSEATSGSGALFASN 299
DB 238 PVLLEFDRDLQRRASFVAVETEMPRLKLLTPMHLVFAA--RGPAPAPGDFA--PVEARR 293
QY 300 VKPGQVYVVLGGGQQLLASHVSVSLREASGAYAPLTAQCTIILNRYLASCYAVIEEH 359
DB 294 LKRGDS--VLAPGGALRPAPARVA--REAVGVFAPLTHGTLAVNDVASCYAVLESH 350
QY 360 SWAHNAFAPFRLAAGLLALCPDGAIPATATTGTGHWTSRLLYRIGSWVL 410
DB 351 QNAHRAFAFLRLIHA-LGALLPGAV-----OPTGHHMYSRLLYRLAEELL 395

Search completed: February 20, 2003, 10:12:37
Job time : 14.0329 secs


```

0Y 1 MVEMLLRLLVFEICALVSSGLTGGPGGJGKRHRPKMLPLAYKOTIPNVAETLG 60
Db 1 MGEMLLRDLGLGFIISLIVPLGLGCGPGJGRRRPOKRLPLAYKOTIPNVEKTLG 60
0Y 61 ASGREYKRTIENSERFELKLPNPNDDIIFKDEENTGADRLMTOCRCKDLNALAISVANOM 120
Db 61 ASGREYVITRNSERFQGLTLPNPNDDIIFKVEENTGADRLMTOCRCKDLNSAISVANOM 120
0Y 121 PGVLRVTEGDEBGGHSEELH EGRNAVDTTSDRBRSKYGMRLAVENAGDPWYTES 180
Db 121 PGVLRVTEGDEGGHSEELH EGRNAVDTTSDRBRSKYGMRLAVENAGDPWYTES 180
0Y 181 KAHIHCSYKAENSYAANKSGCGFPCSATVHLHEHGSTKVLKCLSPGDRVLADADGRLLYS 240
Db 181 KAHIHCSYKAENSYAANKSGCGFPCSATVHLEQGVRIYKGLKPGNRRLAVADVERLLYS 240
0Y 241 FLTEFLDRMOSSRKLPYUETRQPARLLTAHLLTFPAOHPOHNOSEATGSG-----Q 293
Db 241 FLTEFLDEETLRKVEYUETSLPRERRLTAHLLTFPAOEHFGNASGNGNFRSKGHRFR 300
0Y 294 ALFASNVYRPGORVYVLGGGQOOLLPAVSHSVLSREKASGAVAPLTAGCTILINNVLASCY 353
Db 301 SMFASVVRPGIRVVLTEDEREGGLAEATVDRYLL EERTGATAPRTANGTAVIVDRIVLASCY 359
0Y 354 AVIEEHSWAHMAFAPFRLAOGLLAALCP--DGAIPTAATTTGTIHVSRLLYIGISWVL 410
Db 360 AVIEEHSWAHMAFAPRLRGFGILTSFSPDYSSHSHPAPSQAEQVHWYSEILVRIGTWVL 419
0Y 411 DGDMLHPLGNVAPAS 425
Db 420 QADTTHPLGMAAKSS 434

```

RESULT 2			
09W709			
ID	09W709	PRELIMINARY;	PRT: 414 AA.
AC	09W709;		
DT	01-NOV-1999 (TREMBLrel. 12, Created)		
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
DE	Sonic hedgehog.		
GN	SHH.		
OS	Paralichthys olivaceus (Flounder).		
OC	Euteleostei; Metacoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;		
OC	Acanthomorphi; Acanthopterygii; Percomorpha; Pleuronectiformes;		
OC	Pleuronectidae; Paralichthyidae; Paralichthys.		
OX	NCBI_TaxID=8255;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99238226; PubMed=10223710;		
RA	Suzuki T., Ichiro O., Kurikawa T.;		
RT	"Retinoic acid given at late embryonic stage depresses sonic hedgehog		
RT	and Hoxd-4 expression in the pharyngeal area and induces skeletal		
RT	malformation in flounder (Paralichthys olivaceus) embryos.";		
RL	Dev. Growth Differ. 41:143-152(1999).		
DR	EMBL; AB029748; BAA82360.1; -.		
DR	HSSP; 062226; 1VHH.		
DR	MEROPS; C46.001; -.		
DR	InterPro; IPR001767; Hedgehog_hntc.		
DR	InterPro; IPR003586; Hedgehog_hntc.		
DR	InterPro; IPR003587; Hedgehog_hntnc.		
DR	InterPro; IPR000320; HH_signal.		
DR	InterPro; IPR002203; Intein.		
DR	InterPro; IPR001657; SonicHH.		
DR	Pfam; PF01085; HH_signal; 1.		
DR	Pfam; PF01079; Hint; 1.		
DR	PRINTS; PR00632; SONICHHOG.		
DR	ProDom; PD003042; HH_signal; 1.		
DR	SMART; SM00305; Hintc; 1.		
DR	SMART; SM00306; Hintnc; 1.		
SO	SEQUENCE 414 AA; 45945 MW; 50607BF3DB7C0DA3 CRC64;		

Query Match	67.68;	Score 1499;	DB 13;	Length 414;
Best Local Similarity	69.38;	Pred. NO. 3e-115;		
Matches 289;	Conservative 52;	Mismatches 68;	Indels 8;	Gaps 5;

QY	4	MLLTLRLVLWGFI	ALLVSSGLTGC	PRGJRGJGRRRH	PKLTPRAYKOF	IPNVAEKTIGASC	63			
Db	1	MLMTWRIYLA	VICLISLVS	SGMGCGRGY	GRRRH	PKLTPRAYKOF	IPNVAEKTIGASC 60			
QY	64	RYEKGITPNSR	RFELTPNPNY	PDIIIFND	EENTGADRLMTQ	RCKDKLNA	LSVNNQMPGV 1233			
Db	61	RYEKGITPNSR	RFELTPNPNY	TDIIIFND	EENTGADRLMTQ	RCKDKLNA	LSVNNQMPGV 1200			
QY	124	KLRTYEGWDE	DGHSESLH	YEGRAVDIT	SDRDSKYGM	LARLAVAGE	DWYYESKAH 1833			
Db	121	KLRTYEGWDE	DGHHPFES	LHNEGRAVD	ITSDRDKSKY	GTLSRLAVAGE	DWYYESKAH 1800			
QY	184	IHC	SVKAENSVAAK	SGGCGF	PSATVHL	EHGCTKYL	KDLSFEDRYLAADADGRLLYSDFLT 2433			
Db	181	IHC	SVKAENSVAAK	SGGCGF	PSSTVYT	LDQDCKPKVK	MLQGDRLYADADGAPYVYTFDIFM 2400			
QY	244	FLDRNDSR	KFYUET	PROPABRL	LLTAHL	LFAPQHN	OSEATGTSOGALFPA	SVNPKPG 3033		
Db	241	FIDODST	TRKLFYEI	ENDSGO	KITL	RAHLH	LFPG--HNS	TEBAHNGMSAVRASOYRPG 2973		
QY	304	QRYVYL	GEQGOOLL	PASVHSV	LSREESA	GAAPRL	TAAQGITL	INRVLASCYAVIEHSMAH 3633		
Db	298	QTVFVL--	DAERLQ	PVTKRI--	YTOEH	DGSPAPV	TAAQGT	VVDQVLASCAYV	IODHELHAH 3544	
QY	364	WAFAPRL	AGOLLAL	CPDGA	LIPRA	ATTG	TIHMYSK	LLRIGSVW	LDGALPRLGM 4200	
Db	355	WALPAPRL	AAHVVSSLLF--	SSQ	PA	AKOD	CVHMYSK	ILTYOLG	VWLLDHSIH	PLGM 4009

RESULT 3			
057567			
ID	057567	PRELIMINARY;	PRT; 406 AA.
AC	057567;		
DT	01-JUN-1998 (TREMBLrel. 06, Created)		
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Hedgehog segment polarity homolog.		
OS	Notophthalmus viridescens (Eastern newt) (Triturus viridescens)		
OC	Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi		
CC	Notophthalmus.		
CC	NCBI_Taxid=8316;		
CC	[1]		
RP	SEQUENCE FROM N.A.		
RA	Stark D.R., Gates P.B., Brookes J.P., Ferretti P.;		
RT	"Hedgehog homologue from Notophthalmus viridescens.";		
RL	Dev. Dyn. 0:0-(1998).		
CC	-1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE		
CC	PHOSPHORIBOSYLTRANSFERASE FAMILY.		
DR	EMBL; AF047446; AAC03108.1; -.		
DR	HSSP; Q62226; IVH.		
DR	InterPro; IPR001767; Hedgehog_hnt.		
DR	InterPro; IPR003586; Hedgehog_hntC.		
DR	InterPro; IPR003587; Hedgehog_hntN.		
DR	InterPro; IPR000320; HH_signal.		
DR	InterPro; IPR002375; P/Py_TF_transf.		
DR	InterPro; IPR001657; Son1GH.		
DR	Pfam; PF01085; HH_signal; 1.		
DR	Pfam; PF01079; Hntc; 1.		
DR	PRINTS; PRO0632; SON1GHOG.		
DR	ProDom; BD003042; HH_signal; 1.		
DR	SMART; SM00305; Hntc; 1.		
DR	SMART; SM00306; HntN; 1.		
DR	PROSITE; PS00103; PUR_PYR_TRANSF; 1.		
KW	Glycyltransferase; Transferase.		
SO	SEQUENCE 406 AA; 45072 MW; 5842CCAD5314DA00 CRC64;		


```

Db      64 GLENGRITDSERFHTLKNFNTDILFKDEKGTGADRFMTORKDNLALISVNMQMG 123
Oy      123 VKLRVTEGDEGHHSEELHTEGRAVDITTSDBRSKTKMLARLAVEGFDWVYESNA 182
Db      124 VKLRVTEGDEGHHSEELHTEGRAVDITTSDBRSKTKMLARLAVEGFDWVYESNA 183
Oy      183 HHCYSKAEENVAASKSGCFPSATVHLHGGTKLVKDLSPGDRVLADADGRLYSDFL 242
Db      184 HHCYSKAEENVAASKSGCFPSATVHLHGGTKLVKDLSPGDRVLADADGRLYSDFL 243
Oy      243 TELDRDSSRKLFYVETQPARALLTPAHLFPVAPQHNOSATISTGQALFASNVKP 302
Db      244 TELDRDSSRKLFYVETQPARALLTPAHLFPVAPQHNOSATISTGQALFASNVKP 298
Oy      303 GORVYVLGEGGOQLLPASVHSVSLREASGAYAPLTAOGTILINRVLASCAYIEESMA 362
Db      299 GEFLLTPDDGGGFRKVKVYSVTMRRE-KGATAPLTVHGTIVVDVAMSCYALIESQALA 357
Oy      363 HMAFAPRLAOGLLALCPDGAIPR-AAATTTGIIHWSRLYRIGSVLDGALHP 417
Db      358 HMAFAPRLYUOLTSLSLW-DG--PSHDQTLJEGVHWYPSFEYRYGISLVEPTLLHP 410

```

RESULT 6

```

ID      073803 PRELIMINARY: PRT: 442 AA.
AC      073803:
DT      01-AUG-1998 (TREMBLrel. 07, Created)
DT      01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      Fugu hedgehog.
GN      Fugu.
OS      Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC      Acanthomorphi; Acanthopterygii; Perciformes; Tetraodontiformes;
OC      Tetraodontidae; Takifugu.
OX      NCBI_TaxID=31033;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE-99177347; PubMed-10077531;
RA      Gelliner K., Brenner S.;
RT      Analysis of 148 kb of genomic DNA around the wnt1 locus of Fugu
RT      rubripes.
RL      Genome Res. 9:251-258(1999).
DR      EMBL: AF056116; AAC34384.1; -.
DR      HSP: Q62226; 1VNH.
DR      InterPro: IPR001767; Hedgehog_hlnt.
DR      InterPro: IPR003586; Hedgehog_hlntC.
DR      InterPro: IPR003587; Hedgehog_hlntN.
DR      InterPro: IPR000320; HH_signal.
DR      InterPro: IPR002203; Inteln.
DR      Pfam: PF01085; HH_signal.1.
DR      Pfam: PF01079; Hlnt.1.
DR      ProDom: PD003042; HH_signal.1.
DR      SMART: SM00305; HlntC.1.
DR      SMART: SM00306; HlntN.1.
SQ      SEQUENCE 442 AA: 49286 MW: 1CB2BA23B4B748E8 CRC64:

```

Query Match 47.7%; Score 1059; DB 13; Length 442;
 Best Local Similarity 50.5%; Pred. No. 6; 5e-79;
 Matches 223; Conservative 56; Mismatches 115; Indels 48; Gaps 9;

```

Db      11 LVGFICALVSSGLTCGPGRGIGRRHPRKLTPLAYKQIFIPNAKTKGASGRYSGKIT 70
Oy      14 LLAAMSCVWLVOG--CGGPGGIGTRPRKLTAMTYKQFPPLVSNLNGASGRAGKIT 70
Oy      71 RNSERKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALASVNMOPGVKLTVEG 130
Db      71 RNSERKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALASVNMOPGVKLTVEG 130
Oy      131 WEDGHHSEELHTEGRAVDITTSDBRSKTKMLARLAVEGFDWVYESNAHICSVKA 190

```

```

Db      131 WEDGHHSEELHTEGRAVDITTSDBRSKTKMLARLAVEGFDWVYESNAHICSVKA 190
Oy      191 ENSVAASKSGCFPSATVHLHGGTKLVKDLSPGDRVLADADGRLYSDFLTDORMOS 250
Db      191 DMSVAVERKGGCFPSATVHLHGGTKLVKDLSPGDRVLADADGRLYSDFLTDORMOS 250
Oy      251 SRKLFYVETQPARALLTPAHLFPVAPQHNOSATISTGQALFASNVKP 307
Db      251 SRKLFYVETQPARALLTPAHLFPVAPQHNOSATISTGQALFASNVKP 302
Oy      308 VLGEQGOQLLPASVHSVSLREASGAYAPLTAOGTILINRVLASCAYIEESMAHAP 367
Db      303 VHTAGG-EVHPKRVISIT-EESVSGAYAPLTEGVSFVGVGLASSALVEDIQLAHMAG 360
Oy      368 PFRLL-----AOGLLALCPDGAIPR-AAATTTGIIHWSRLYRIGSVLDGALHP 395
Db      361 PVRLLSSVQOLMAPEEESDSKTPLOPHALVGRDKKCARNSISVREAPRGRTSEV 420
Oy      396 HWSRLYRIGSVLDGALHP 417
Db      421 HMYAQLHRLGWTIVLPDLFHP 442

```

RESULT 7

```

ID      09GRAB PRELIMINARY: PRT: 447 AA.
AC      09GRAB:
DT      01-MAR-2001 (TREMBLrel. 16, Created)
DT      01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      Hedgehog protein (Fragment).
OS      Gryllus bimaculatus (Two-spotted cricket).
OC      Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC      Gryllinae; Gryllus.
OX      NCBI_TaxID=6999;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE-20461145; PubMed-11003837;
RA      Nira N., Inoue Y., Nozawa A., Saito M., Misumi Y., Ohuchi H.,
RA      Yoshioke H., Noji S.;
RT      Correlation of diversity of leg morphology in Gryllus bimaculatus
RT      (cricket) with divergence in dpp expression pattern during leg
RT      development.
RL      Development 127:4373-4381(2000).
DR      EMBL: AB044709; BAB19658.1; -.
DR      HSP: Q62226; 1VNH.
DR      InterPro: IPR001767; Hedgehog_hlnt.
DR      InterPro: IPR003586; Hedgehog_hlntC.
DR      InterPro: IPR003587; Hedgehog_hlntN.
DR      InterPro: IPR000320; HH_signal.
DR      InterPro: IPR001657; Scn1cHH.
DR      Pfam: PF01085; HH_signal.1.
DR      Pfam: PF01079; Hlnt.1.
DR      PRINTS: PR00632; SONICHHOG.
DR      ProDom: PD003042; HH_signal.1.
DR      SMART: SM00305; HlntC.1.
DR      SMART: SM00306; HlntN.1.
FT      NON_TER
SQ      SEQUENCE 447 AA: 48008 MW: 29AEPR061C3EB6F0 CRC64:

```

Query Match 47.1%; Score 1044.5; DB 5; Length 447;
 Best Local Similarity 53.9%; Pred. No. 1e-77;
 Matches 220; Conservative 48; Mismatches 123; Indels 17; Gaps 8;

```

Oy      3 EMLLTRILLVGFICALVSSGLTCGPGRGIGRRHPRKLTPLAYKQIFIPNAEKTGAS 62
Db      14 KPIILLRMLL---LLLMGGAARACGPGRGIGRRASRLTLPVFKQHPVNSERTLGAA 70
Oy      63 GREGKGTNSRFRKLTTPYNPDIIFKDEENTGADRLMTORCKDKLNALASVNMQMG 122
Db      71 GPARVAVDDPRFDLVPNYNADIVFKDEEGTADRLMTORCKEKLNTLALISVNMQMG 130

```


OS Ambystoma mexicanum (Axolotl).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandridae; Ambystomatidae;
 OC Ambystoma.
 OX NCBI_TaxID=8296;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Torok M.A., Izpielua-Belmonte J.C., Gardiner D.M., Bryant S.V.;
 RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF031480; AAD18128.1; -
 DR HSSP; 062226; LVHH.
 DR InterPro: IPR0011767; Hedgehog_hnt.
 DR InterPro: IPR000320; HH_signal.
 DR InterPro: IPR001657; SonicHH.
 DR Pfam: PF01085; HH_signal; 1.
 DR Pfam: PF01079; Hnt; 1.
 DR PRINTS: PR00632; SONICHHOG.
 DR ProDom: PD003042; HH_signal; 1.
 FT NON_TER 1
 FT NON_TER 150
 SQ SEQUENCE 150 AA; 16599 MW; 986F65037A69A8AC CRC64;

Query Match 33.2%; Score 737; DB 13; Length 150;
 Best Local Similarity 93.3%; Pred. No. 4, 5e-53;
 Matches 140; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 86 DIFKDEMTGADRLMTQCKDKLNALAI SVNMQMPGVKLRVTEGWDGHHSESLHAYE 145
 DB 1 DIFKDEMTGADRLMTQCKDKLNALAI SVNMQMPGVKLRVTEGWDGHHSESLHAYE 60
 OY 146 GRAVDITTSDDRSKYGMLARLAVENGFDWYVESKAHHCYSKAVNSAANGCPEPS 205
 DB 61 GRAVDITTSDDRSKYGMLARLAVENGFDWYVESKAHHCYSKAVNSAANGCPEPS 120
 OY 206 ATVLEHGKTKLVKDLSPGDRVLADADGR 235
 DB 121 AKVLEHGVTTPKDLRPGDRVLADADGR 150

RESULT 11

ID 09XS16 PRELIMINARY; PRT; 139 AA.
 AC 09XS16;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
 DE Sonic hedgehog (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=TOOTH GERM;
 RA Koyama E., Iwamoto M., Ohmori T., Kurisu K., Wu C., Ookura T.,
 RA Basalir M.M., Tucker T., Pacifici M.;
 RT "Development of Stratum Intermedium and its Role as a Sonic Hedgehog-
 RT Signaling Structure During Odontogenesis."
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF144100; AAD33926.1; -
 DR HSSP; 062226; LVHH.
 DR InterPro: IPR000320; HH_signal.
 DR InterPro: IPR001657; SonicHH.
 DR Pfam: PF01085; HH_signal; 1.
 DR PRINTS: PR00632; SONICHHOG.
 DR ProDom: PD003042; HH_signal; 1.
 FT NON_TER 1
 FT NON_TER 139
 SQ SEQUENCE 139 AA; 15961 MW; C83B0B96982046B1 CRC64;

Query Match 33.1%; Score 735; DB 6; Length 139;
 Best Local Similarity 99.3%; Pred. No. 5, 8e-53;

Matches 138; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 46 AKYQFIPNVAEKTIGASGRYEGKITRNSERFKELTPNYNPDIIFKDEMTGADRLMTQCK 105
 DB 1 AKYQFIPNVAEKTIGASGRYEGKITRNSERFKELTPNYNPDIIFKDEMTGADRLMTQCK 60
 OY 106 KDKLNALAI SVNMQMPGVKLRVTEGWDGHHSESLHAYEGRAVDITTSDDRSKYGMLA 165
 DB 61 KDKLNALAI SVNMQMPGVKLRVTEGWDGHHSESLHAYEGRAVDITTSDDRSKYGMLA 120
 OY 166 RLAVENGFDWYVESKAH 184
 DB 121 RLAVENGFDWYVESKAH 139

RESULT 12

ID 09W6C1 PRELIMINARY; PRT; 138 AA.
 AC 09W6C1;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
 DE Sonic hedgehog protein (Fragment).
 GN SHH.
 OS Eleutherodactylus coqui (Puerto Rican coqui).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Leptodactylidae;
 OC Eleutherodactylus.
 OX NCBI_TaxID=57060;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Carl T.F., Richardson M.K., Olsson L., Schlosser G., Klimowsky M.W.,
 RA Hanken J.;
 RT "Differences in vertebrate limb development revealed by studies of the
 RT direct developing frog E. coqui".
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF113403; AAD23436.1; -
 DR HSSP; 062226; LVHH.
 DR InterPro: IPR000320; HH_signal.
 DR InterPro: IPR001657; SonicHH.
 DR Pfam: PF01085; HH_signal; 1.
 DR PRINTS: PR00632; SONICHHOG.
 DR ProDom: PD003042; HH_signal; 1.
 FT NON_TER 1
 FT NON_TER 138
 SQ SEQUENCE 138 AA; 15751 MW; FFA156A1F4681F0 CRC64;

Query Match 32.1%; Score 713; DB 13; Length 138;
 Best Local Similarity 97.1%; Pred. No. 3, 8e-51;
 Matches 134; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 51 IPNVAEKTIGASGRYEGKITRNSERFKELTPNYNPDIIFKDEMTGADRLMTQCKDKLN 110
 DB 1 IPNVAEKTIGASGRYEGKITRNSERFKELTPNYNPDIIFKDEMTGADRLMTQCKDKLN 60
 OY 111 ALAISVNMQMPGVKLRVTEGWDGHHSESLHAYEGRAVDITTSDDRSKYGMLARLAVE 170
 DB 61 ALAISVNMQMPGVKLRVTEGWDGHHSESLHAYEGRAVDITTSDDRSKYGMLARLAVE 120
 OY 171 AGFDWYVESKAHHCYSV 188
 DB 121 AGFDWYVESKAHHCYSV 138

RESULT 13

ID 09WV29 PRELIMINARY; PRT; 177 AA.
 AC 09WV29;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
 DE Indian hedgehog protein (Fragment).
 OS Rattus norvegicus (Rat).

```

OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RA Garges P.L., Meyer R.A. Jr., Brown C.A., Price D.K.;
RT "Indigen hedgehog in rat.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF162914; AAD45372.1; -
DR HSSP: Q62226; 1VHH.
DR MEROPS: C46_003; -
DR InterPro: IPR001767; Hedgehog_hlnt.
DR InterPro: IPR003587; Hedgehog_hlntn.
DR InterPro: IPR00320; HH_signal.
DR InterPro: IPR001657; SonICHH.
DR Pfam: PF01085; HH_signal; 1.
DR Pfam: PF01079; Hlnt; 1.
DR PRINTS: PRO0632; SONICHOG.
DR PRODOM: PD003042; HH_signal; 1.
DR SMART: SM00306; Hlntn; 1.
FT NON_TER 1
FT SEQUENCE 177 AA; 19739 MW; CA5871626A56565 CRC64;

Query Match
Best Local Similarity 31.4%; Score 697; DB 11; Length 177;
Matches 132; Conservative 14; Mismatches 31; Indels 0; Gaps 0;

QY 100 LMTORCKDLNALAISVNNOMPQVKLRTGEMWDEGHSESLHTEGRAVDTTSDRRS 159
DB 1 LMTORCKDLRLSLAISVNNOMPQVKLRTGEMWDEGHSESLHTEGRAVDTTSDRRN 60
QY 160 KYGMALARLAVAGEFDWYTESKAHHCSSVKAENSVAAKSGCFPSATVHLEHGTRLYK 219
DB 61 KYGLLARLAVAGFDWYTESKAHHCSSVKAENSVAAKSGCFPSAGAVHLEHGERVALS 120
QY 220 DLSFGDRLVADADGRLLYSDFLFLDRMDSRKLYVIETRQPRARLLTAHLLF 276
DB 121 AVKGDRLVLAGDEGDNPTFSVLLFLDRPRRLRAFYIETQPPRRLTAHLLF 177

RESULT 14
ID 042234 PRELIMINARY; PRT; 137 AA.
AC 042234;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Sonic hedgehog (fragment).
GN SHH.
OS Cocurnix coturnix (common quail).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Cocurnix.
OC NCBI_TaxID=9091;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE-98167903; PubMed-9435297;
RA Borycki A.G., Mendham L., Emerson C.P. Jr.;
RT "Control of somite patterning by sonic hedgehog and its downstream
signal response genes.";
RL Development 125:777-790(1998).
DR EMBL: AF022882; AAB80949.1; -
DR MEROPS: C46_002; -
DR InterPro: IPR001767; Hedgehog_hlnt.
DR InterPro: IPR003586; Hedgehog_hlntc.
DR InterPro: IPR001657; SonICHH.
DR Pfam: PF01079; Hlnt; 1.
DR PRINTS: PRO0632; SONICHOG.
DR SMART: SM00305; Hlntc; 1.
FT NON_TER 1
FT SEQUENCE 137 AA; 14571 MW; 7FD29DF815AF1532 CRC64;

```

```

Query Match
Best Local Similarity 30.7%; Score 682; DB 13; Length 137;
Matches 130; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 290 TSGCALFASNVKPGORVYVLEGGQQLPASVSHSLREESGAVAPLTAOCTILINRVL 349
DB 2 TSGCAFFASNVKPGORVYVLEGGROLPLASVSHSLREESGAVAPLTAOCTILINRVL 61
QY 350 ASCYAVIEHSMWAFPRLAGCLALCPDGAIPATTTGIIHYSRLYRIGSV 409
DB 62 TSCTAVIEHSMWAFPRLAGCLALCPDGAIPATTTGIIHYSRLYRIGSV 121
QY 410 LDGDALHPLGNVAPAS 425
DB 122 LDGDALHPLGNVAPAS 137

RESULT 15
ID 096699 PRELIMINARY; PRT; 185 AA.
AC 096699;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Hedgehog protein (fragment).
GN HH.
OS Junonia coenia (Peacock butterfly) (Precis coenia).
OC Eukaryota; Metazoa; Archopoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditylsia;
OC Papilionoidea; Nymphalidae; Nymphalinae; Junonia.
OC NCBI_TaxID=39708;
RN 11
RP SEQUENCE FROM N.A.
RA Keys D.N., Lewis D.L., Selegue J.E., Pearson B.J., Goodrich L.V.,
RA Johnson R.L., Gates J., Scott M.P., Carroll S.B.;
RT "Recruitment of a hedgehog regulatory circuit in butterfly eyespot
evolution.";
RL Science 0:0-0(1999).
DR EMBL: AF117742; AAD08931.1; -
DR HSSP: Q62226; 1VHH.
DR InterPro: IPR00320; HH_signal.
DR InterPro: IPR001657; SonICHH.
DR Pfam: PF01085; HH_signal; 1.
DR PRINTS: PRO0632; SONICHOG.
DR PRODOM: PD003042; HH_signal; 1.
FT NON_TER 1
FT SEQUENCE 185 AA; 20745 MW; E9B9F4CE2FAAE662 CRC64;

Query Match
Best Local Similarity 28.1%; Score 624; DB 5; Length 185;
Matches 117; Conservative 21; Mismatches 30; Indels 2; Gaps 2;

QY 61 ASGREGITRNSERKELTPYNPDIIFKDEENTGADRLMTORCKDLNALAISVNNOW 120
DB 3 ASGPREGITRDEKFRDLVPYNPDIIDFKDEGTGADRLMTORCKELNTLAISVNNOW 62
QY 121 PGVKLRVTEGWDGHHSESLHTEGRAVDTTSDRRSKKGMALARLAVAGFDWYTES 180
DB 63 PGVRLRVTEGWDGHHSESLHTEGRAVDTTSDRRSKKGMALARLAVAGFDWYTES 122
QY 181 KAHHCSSVKAENSVAAKSGCFPSATVHLEHGTRLYKDLSPGRLVLA 230
DB 123 RSYHCSSVKTSSVGT-GAGCFPSGAVYHTENGPRD-IASLKKKGVLA 170

```

Search completed: February 20, 2003, 10:11:50
Job time : 30.337 secs

THIS PAGE BLANK (USPTO)

PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
XX Ingham PW, McMahon AP, Tabin CJ;
XX WPI: 1995-255060/33.
XX N-PSDB; AA091637.
XX Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
PT to treat degenerative nervous system disorder(s) and in gene
PT therapy.
XX Claim 17; Page 139-41; 210pp; English.
XX
XX The sequence represents a mouse sonic hedgehog protein,
CC homologous to a Drosophila hedgehog protein (AA077337), and is
CC encoded by a cDNA isolated from an 8.5-day post coitum mouse cDNA
CC library. Probes and primers derived from the sonic hedgehog gene
CC may be used as diagnostic agents for neuromuscular, autonomic or
CC central nervous system disorders, and the gene may also be used in
CC gene therapy. Antibodies generated from the protein may be used
CC as therapeutic or research reagents.

XX Sequence 437 AA:

Query Match 100.0%; Score 2289; DB 16; Length 437;
Best Local Similarity 100.0%; Pred. No. 9,4e-226;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLLLLARCFVIYIASSLLVCPGLACGPGRGFRGRRHPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVIYIASSLLVCPGLACGPGRGFRGRRHPKLTPLAYKQFIPNVAEKTIGAS 60
OY 61 GREGKITSRSEKFKELTPYNYNDIIFKDEBENTGADRLMTORCKDKINALAISVMNQWPG 120
DB 61 GREGKITSRSEKFKELTPYNYNDIIFKDEBENTGADRLMTORCKDKINALAISVMNQWPG 120
OY 121 VRLRVTEGMEDEGHSESLHYEGRAVDITTSRDNSKYGMALARAVEAGFDWVYYESKA 180
DB 121 VRLRVTEGMEDEGHSESLHYEGRAVDITTSRDNSKYGMALARAVEAGFDWVYYESKA 180
OY 181 HICSVAKENSVAAGSGCGPGSATVHLEOGGKFLVLDLRPGDRLVAADOGRLYSDFL 240
DB 181 HICSVAKENSVAAGSGCGPGSATVHLEOGGKFLVLDLRPGDRLVAADOGRLYSDFL 240
OY 241 TFLDRDEGAARKVYVITLPERRELLTAAHLLEFVAPHNDSGPTPGSALFASVRPGQR 300
DB 241 TFLDRDEGAARKVYVITLPERRELLTAAHLLEFVAPHNDSGPTPGSALFASVRPGQR 300
OY 301 VYVVAERGGRRLLPRAVHVTLREERAGAVAPLTAGTILINRVLASCYAVEEHSMAH 360
DB 301 VYVVAERGGRRLLPRAVHVTLREERAGAVAPLTAGTILINRVLASCYAVEEHSMAH 360
OY 361 RAFAFRLIALALLAALPARPTDGGGGSIPAAOSATEARGAEPAGIHWYSOLLHYHGTW 420
DB 361 RAFAFRLIALALLAALPARPTDGGGGSIPAAOSATEARGAEPAGIHWYSOLLHYHGTW 420
OY 421 LLDSETHMPLGMAYKSS 437
DB 421 LLDSETHMPLGMAYKSS 437

RESULT 2

AA05857 standard; Protein: 437 AA.

AA05857;

02-AUG-1999 (first entry)

Mouse Sonic hedgehog Shh protein.

Sonic hedgehog; Shh; mouse; epithelial tissue; epithelium;
cutaneous tissue; skin; hair; wound healing; vulvectomy;

KW burn; skin grafting; pressure sore; ulcer; ulcerative colitis;
KW alopecia; psoriasis; keratosis; acne; comedogenic lesion;
KW folliculitis; pseudofolliculitis; keratoacanthoma; callosities;
KW Darier's disease; scar; autoimmune disease; pemphigus;
KW epidermolysis; lupus lesion; desquamative lesion; carcinoma;
KW therapy; hedgehog therapeutic; ptc therapeutic; patched.
XX Mus musculus.
XX WO920298-A1.
XX 29-APR-1999.
XX 20-OCT-1998; 98WO-US22227.
XX 11-SEP-1998; 98US-0151999.
XX 20-OCT-1997; 97US-0955552.
XX (ONTO-) ONTOGENY INC.
XX Wang EA;
XX WPI: 1999-288170/24.
XX N-PSDB; AAX25620.
XX Use of hedgehog polypeptides on patched therapeutics
PS Claim 26; Page 124-125; 146pp; English.

XX The present sequence represents mouse Sonic hedgehog protein
CC Shh. The invention relates to a method for modulating the growth
CC state an epithelial cell by ecotopically contacting the epithelial
CC cell, in vitro or in vivo, with a hedgehog therapeutic (i.e. a
CC hedgehog polypeptide or gene therapy construct) or ptc therapeutic
CC (i.e. a small organic molecule that mimics the effect of hedgehog
CC proteins on patched signaling, or activates or potentiates patched
CC signaling) in an amount effective to alter the rate of proliferation
CC comprises at least a bioactive extracellular portion of a hedgehog
CC protein (see AA05854-62) encoded by a vertebrate hedgehog gene (see
CC AAX25617-25), especially a human hedgehog gene. Promotion of
CC proliferation of epithelial cells can be used to control a wound
CC healing process in e.g. burn treatment, skin regeneration, skin
CC grafting, pressure sore treatment, dermal ulcer treatment, post
CC surgery scar reduction or treatment of ulcerative colitis (claimed).
CC It can also be used to induce hair growth for the treatment of
CC alopecia (claimed). Inhibition of the growth of epithelial tissue
CC can be used to treat or prevent hyperplastic or neoplastic
CC conditions, e.g. psoriasis, keratosis, acne, comedogenic lesions,
CC folliculitis and pseudofolliculitis, keratoacanthoma, callosities,
CC Darier's disease, keloids, hypertrophic scars, or autoimmune
CC disorders, e.g. aphthous ulcers, pemphigus vulgaris, pemphigus
CC foliaceus, pemphigus vegetans, pemphigus erythematous,
CC epidermolysis, lupus lesions, desquamative lesions or carcinomas.
CC The methods can also be used to counteract the effects of ageing on
CC skin.

SO Sequence 437 AA:

Query Match 100.0%; Score 2289; DB 20; Length 437;
Best Local Similarity 100.0%; Pred. No. 9,4e-226;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLLLLARCFVIYIASSLLVCPGLACGPGRGFRGRRHPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVIYIASSLLVCPGLACGPGRGFRGRRHPKLTPLAYKQFIPNVAEKTIGAS 60
OY 61 GREGKITSRSEKFKELTPYNYNDIIFKDEBENTGADRLMTORCKDKINALAISVMNQWPG 120
DB 61 GREGKITSRSEKFKELTPYNYNDIIFKDEBENTGADRLMTORCKDKINALAISVMNQWPG 120
OY 121 VRLRVTEGMEDEGHSESLHYEGRAVDITTSRDNSKYGMALARAVEAGFDWVYYESKA 180
DB 121 VRLRVTEGMEDEGHSESLHYEGRAVDITTSRDNSKYGMALARAVEAGFDWVYYESKA 180

Db	121	VLRLVTEGWDGDEHHSEBSLHYEGRAVDITTSDBRONSKYOMLARLAVAEAGFDWYYIESKA	180
QY	181	HIHCSVKAENSVAAKSGGCEPFSATVHLEOGGTLKYVDRLPBGRVYLAADOGRLTSDFL	240
Db	181	HIHCSVKAENSVAAKSGGCEPFSATVHLEOGGTLKYVDRLPBGRVYLAADOGRLTSDFL	240
QY	241	TFLDORDEAKKVFVETLEPRELLTLTAHLTPVARNHNSGPPGSAFASRVPRGQR	300
Db	241	TFLDORDEAKKVFVETLEPRELLTLTAHLTPVARNHNSGPPGSAFASRVPRGQR	300
QY	301	VYVVAERGGDRRLPAVAHVSTLLEEBAAGVAPLTAGTILLINRVLASCYAVIEESHMAH	360
Db	301	VYVVAERGGDRRLPAVAHVSTLLEEBAAGVAPLTAGTILLINRVLASCYAVIEESHMAH	360
QY	361	RAFAPEFLAHLAALPARFDGGGGSIPAAOSATBARCAEAPAGIHWMSOLLYHIGTW	420
Db	361	RAFAPEFLAHLAALPARFDGGGGSIPAAOSATBARCAEAPAGIHWMSOLLYHIGTW	420
QY	421	LIDSETMHPLGMAVKSS	437
Db	421	LIDSETMHPLGMAVKSS	437

RESULT 3
 ID AAI05513
 AAY05513 standard; Protein; 437 AA.
 XX
 AC
 AAY05513;
 DT 05-JUL-1999 (first entry)
 XX
 DE Mouse Sonic hedgehog protein Shh.
 XX
 KW Sonic hedgehog; Shh protein; mouse; hedgehog therapeutic;
 KW ptc therapeutic; patched; signal transduction; muscle atrophy;
 KW cachexia; muscular myopathy; myoblastic sarcoma; therapy.
 XX
 OS Mus sp.
 XX
 PN MO9910004-A2.
 PD 04-MAR-1999.
 XX
 PR 28-AUG-1998; 98WO-US17922.
 XX
 PR 29-AUG-1997; 97US-0057394.
 XX
 PA (ONTO-) ONTOGENY INC.
 XX
 PI Bladgen CS, Currie PD, Hughes SM, Ingham PW;
 DR WPI; 1999-243557/20.
 DR N-PSDB; AAX25101.
 XX
 PT A new method to regulate muscle growth
 XX
 PS Disclosure; Page 115-116; 130pp; English.
 CC
 CC The present sequence is mouse Sonic hedgehog protein Shh. The
 CC invention relates to a method for modulating the formation and/or
 CC maintenance of muscle tissue by ecotopically contacting muscle
 CC cells, especially muscle stem/progenitor cells, in vitro or in
 CC vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and
 CC gene therapy constructs) or ptc therapeutic (i.e. a small organic
 CC molecule that mimics the effect of hedgehog proteins on patched
 CC signalling, or activates or potentiates patched signalling) in an
 CC amount effective to alter the growth state of the treated cells.
 CC Also claimed is a method for treatment or prevention of disorders
 CC of, or surgical or cosmetic repair of, such muscle tissues, by
 CC administering a hedgehog polypeptide or ptc therapeutic. The
 CC disorder may be muscle atrophy, in particular skeletal muscle
 CC atrophy or cardiac muscle atrophy, cachexia, or muscular myopathy
 CC (all claimed). The hedgehog polypeptide or ptc therapeutic can

CC inhibit growth of myoblastic-derived tissue to provide treatment of
CC hyperplastic or neoplastic growth of muscle tissue such as in
CC myoblastic sarcoma (also claimed). The hedgehog therapeutic
CC preferably comprises at least a bioactive extracellular portion of
CC a hedgehog protein (see A10510-19) encoded by a vertebrate hedgehog
CC gene (see AA25098-107), especially a human hedgehog gene.

[illegible]

Oy 1 MLLTLARCEFLVIASSLVCYGLACGPGRGGRGRHPKLTPLAYKQIIPNVAETIGAS 60

Db 1 MLLTLARCEFLVIASSLVCYGLACGPGRGGRGRHPKLTPLAYKQIIPNVAETIGAS 60

Oy 61 GRYECKITRNSEREKELTPNYPNDIIEFDEBENTGADRLMTORCKDKLNALAI SVYNOMP 120

Db 61 GRYECKITRNSEREKELTPNYPNDIIEFDEBENTGADRLMTORCKDKLNALAI SVYNOMP 120

OY	121	VLRLVTEBMDEGHHSESLIIEGKAVAVITTSDBDRSKYGMRLARLAVEAGFDWVYESKA	180
OY	121	VLRLVTEBMDEGHHSESLIIEGKAVAVITTSDBDRSKYGMRLARLAVEAGFDWVYESKA	180
Db	121	VLRLVTEBMDEGHHSESLIIEGKAVAVITTSDBDRSKYGMRLARLAVEAGFDWVYESKA	180
OY	181	HLHCSVKAEKNSVAAKSGGCGPFGSATVHLDEOGCTKLVKDLPKPGDRVLAADDGRLYSDFL	240
Db	181	HLHCSVKAEKNSVAAKSGGCGPFGSATVHLDEOGCTKLVKDLPKPGDRVLAADDGRLYSDFL	240
OY	241	TFLLDDEGAKKVFYIETLEPERELLTLAHLLEVAHPNDSGPTPGPSALFASVRRPQR	300
Db	241	TFLLDDEGAKKVFYIETLEPERELLTLAHLLEVAHPNDSGPTPGPSALFASVRRPQR	300
OY	301	VYVVAERGGDRRLPLAAVHSVTLREEAGAVAPLTAHGTILINRYLASCAYVIEHSMW	360
Db	301	VYVVAERGGDRRLPLAAVHSVTLREEAGAVAPLTAHGTILINRYLASCAYVIEHSMW	360
OY	361	RAFAFPRLAHALLAALAPARTDGGGGGSIIPAQSAATEARGAEPPTAGIHWSQLLYHIGTW	420
Db	361	RAFAFPRLAHALLAALAPARTDGGGGGSIIPAQSAATEARGAEPPTAGIHWSQLLYHIGTW	420
OY	421	LIDSETMHPILGMVAKSS 437	
Db	421	LIDSETMHPILGMVAKSS 437	
RESULT 4			
ID	AAW97768	AAW97768 standard; Protein: 437 AA.	
XX	AAW97768;		
AC			
XX			
DT	21-MAY-1999	(first entry)	
XX			
DE	Mouse Sonic hedgehog (Shh) protein.		
XX			
KW	Sonic hedgehog; Shh protein; mouse; dopaminergic; GABA-nergic;		
KW	PtC therapeutic; patched; signal transduction; Parkinson's disease;		
KW	Huntington's disease; amyotrophic lateral sclerosis;		
KW	cerebral ischaemia; hypoxia; neuroprotective; therapy.		
XX			
OS	Mus sp.		
XX			
PN	W09904775-A2.		
XX			
PD	04-FEB-1999.		
XX			
PF	24-JUL-1998;	98WO-US15419.	
XX			
PR	24-JUL-1997;	97US-0900220.	
XX			
PA	(ONTO-) ONTOGENY INC.		
XX			

PI Mahanthappa NK, Miao N, Pang K, Wang M;
XX WPI; 1999-142578/12.
DR N-PSDB; AAX07274.
XX
PT Increasing the survival of neuronal, dopaminergic and GABA-nergic
PT cells - by using a ptc therapeutic such as a protein kinase
PT inhibitor, or an agent derived from hedgehog polypeptides, useful in
PT the treatment of Parkinson's disease
XX
PS Disclosure; Page 89-91; 138pp; English.
XX
CC This polypeptide is mouse Shh Sonic hedgehog protein. The
CC invention is based on the finding that hedgehog proteins are useful
CC as protective agents in the treatment and prophylaxis of
CC neurodegenerative disorders resulting from the loss of dopaminergic
CC and/or GABA-nergic neurons, or the general loss of tissue from the
CC substantia nigra. Exemplary disorders include Parkinson's disease,
CC Huntington's disease (both claimed), amyotrophic lateral sclerosis
CC and cerebral ischemia. The invention relates to hedgehog
CC therapeutics (i.e. hedgehog polypeptides and gene therapy constructs
CC e.g. constructs encoding recombinant hedgehog polypeptides and
CC trans-activation constructs for altering hedgehog gene regulatory
CC sequences) and ptc therapeutics (i.e. agents which mimic the effect
CC of naturally occurring hedgehog proteins on patched signaling)
CC and Dh polypeptides (see AAW97763-64) are preferred. Human Ihh
CC can also be used for the maintenance of differentiated neurons in
CC cultures, and to enhance the implantation of such neuronal cells in
CC an animal. They can be used to prevent or treat neurodegenerative
CC conditions arising from the use of certain drugs, and in the
CC prevention and/or treatment of hypoxia, e.g. as a neuroprotective
CC agent.
XX
SQ Sequence 437 AA;
Query Match 100.0%; Score 2289; DB 20; Length 437;
Best Local Similarity 100.0%; Pred. No. 9.4e-226;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLLILARCFVLIASSLVCPGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTGAS 60
DB 1 MLLILARCFVLIASSLVCPGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTGAS 60
QY 61 GRYEGKITRNSERKELTPYNNPDIIFFKDEENTGADRLMTORCKDKINALAISVMNMPG 120
DB 61 GRYEGKITRNSERKELTPYNNPDIIFFKDEENTGADRLMTORCKDKINALAISVMNMPG 120
QY 121 VRLVTEGMDDEGHSESLHYEGRAVDITTSRDRSKYGLARLAVEAGFDWYVESKA 180
DB 121 VRLVTEGMDDEGHSESLHYEGRAVDITTSRDRSKYGLARLAVEAGFDWYVESKA 180
QY 181 HHCYSKAENSVAAKSGCGPGSATVHLEGGGKLVKDLRPGGRVLAADOGGLLSDFL 240
DB 181 HHCYSKAENSVAAKSGCGPGSATVHLEGGGKLVKDLRPGGRVLAADOGGLLSDFL 240
QY 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLLEVAHPNDSGPTPGSALFASRVAPGGR 300
DB 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLLEVAHPNDSGPTPGSALFASRVAPGGR 300
QY 301 VYVVAERGDRRLPAAVHSVTLREEEAGAVAPLTAHGILLNRYLASCYAVIEEHSMAH 360
DB 301 VYVVAERGDRRLPAAVHSVTLREEEAGAVAPLTAHGILLNRYLASCYAVIEEHSMAH 360
QY 361 RAFAFRLAHLALALAPARTDGGGGSIPAAQSATEARGAEPAGIHWYSOLLHIGTW 420
DB 361 RAFAFRLAHLALALAPARTDGGGGSIPAAQSATEARGAEPAGIHWYSOLLHIGTW 420
QY 421 LLDSTHNPILGMAVSS 437
DB 421 LLDSTHNPILGMAVSS 437

RESULT 5
AAW94471
ID AAW94471 standard; Protein; 437 AA.
XX
AC AAW94471;
XX
DT 29-APR-1999 (first entry)
XX
DE Mouse Shh hedgehog protein sequence.
XX
KW Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;
KW brain infarction; cerebral infarction; transient ischemic attack;
KW stroke; cerebral infarct volume; spinal cord; edema; trauma;
KW haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.
OS Mus sp.
XX
PN WO9900117-A2.
XX
PD 07-JAN-1999.
XX
PF 26-JUN-1998; 98WO-US13387.
XX
PR 27-JUN-1997; 97US-0883656.
XX
PA (ONTO-) ONTOGENY INC.
XX
PI Mahanthappa NK;
XX
DR WPI; 1999-095458/08.
DR N-PSDB; AAX16185.
XX
PT Method for limiting damage to neurons caused by ischemic or epoxic
PT conditions - is used for the treatment and prevention of e.g.
PT cerebral infarction, stroke and transient ischemic attacks
XX
PS Disclosure; Page 68-70; 104pp; English.
XX
CC A method has been developed for limiting the damage to neuronal cells by
CC ischemic or epoxic conditions by administering a ptc (patched)
CC therapeutic agent to reduce cerebral infarct volume (CIIV). Damage to
CC neuronal cells can also be limited by administering a gene activation
CC construct which recombines with the genomic hedgehog gene to provide a
CC heterologous transcription regulator linked to the coding region of this
CC gene. Administration of the ptc therapeutic agent is used to protect
CC cerebral tissues against ischemic injury; to treat cerebral infarct or
CC ischemia; stroke (thrombotic or embolic) and transient ischemic
CC attacks. It may also be used as a prophylactic in many other cases of
CC injury to the brain or spinal cord, edema caused by trauma, haemorrhage
CC and encephalomyelitis, or in conjunction with (coronary bypass) surgery.
CC treatment (which may be prophylactic) is used where ischemic/epoxic
CC conditions may cause cerebral hypoxia, or progressive loss of neurons
CC due to oxygen depletion, including in patients with hypotension. The
CC present sequence represents a hedgehog sequence given in the present
CC invention.
XX
SQ Sequence 437 AA;
Query Match 100.0%; Score 2289; DB 20; Length 437;
Best Local Similarity 100.0%; Pred. No. 9.4e-226;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLLILARCFVLIASSLVCPGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTGAS 60
DB 1 MLLILARCFVLIASSLVCPGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTGAS 60
QY 61 GRYEGKITRNSERKELTPYNNPDIIFFKDEENTGADRLMTORCKDKINALAISVMNMPG 120
DB 61 GRYEGKITRNSERKELTPYNNPDIIFFKDEENTGADRLMTORCKDKINALAISVMNMPG 120
QY 121 VRLVTEGMDDEGHSESLHYEGRAVDITTSRDRSKYGLARLAVEAGFDWYVESKA 180
DB 121 VRLVTEGMDDEGHSESLHYEGRAVDITTSRDRSKYGLARLAVEAGFDWYVESKA 180

Db 121 VRLNRTGMDGHHSESLHYEGRAVDITTSDRRSKYGMARLAVEAGFDWYYESKA 180
Qy 181 HHC5VKAENSVAAKSGCGCPGSAATVHLEGGCTKLVKDLRPGDRVLAADGGRLLYSDFL 240
Cc 181 HHC5VKAENSVAAKSGCGCPGSAATVHLEGGCTKLVKDLRPGDRVLAADGGRLLYSDFL 240
Db 181 HHC5VKAENSVAAKSGCGCPGSAATVHLEGGCTKLVKDLRPGDRVLAADGGRLLYSDFL 240
Qy 241 TFLDRDEGAKKVFYIETLEPRERLLTAHLLFVAPHNDGSPPPGSPALFASRVPRQOR 300
Cc 241 TFLDRDEGAKKVFYIETLEPRERLLTAHLLFVAPHNDGSPPPGSPALFASRVPRQOR 300
Db 241 TFLDRDEGAKKVFYIETLEPRERLLTAHLLFVAPHNDGSPPPGSPALFASRVPRQOR 300
Qy 301 VYVAERGGDRRLPAAVHSVTLREBEAGAYAPLTAHGTILLINRYLASCYAVIEESHNAH 360
Cc 301 VYVAERGGDRRLPAAVHSVTLREBEAGAYAPLTAHGTILLINRYLASCYAVIEESHNAH 360
Db 301 VYVAERGGDRRLPAAVHSVTLREBEAGAYAPLTAHGTILLINRYLASCYAVIEESHNAH 360
Qy 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPRTAGIHMYSQLYHIGTW 420
Cc 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPRTAGIHMYSQLYHIGTW 420
Db 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPRTAGIHMYSQLYHIGTW 420
Qy 421 LLDSETHPLGMAVKSS 437
Cc 421 LLDSETHPLGMAVKSS 437
Db 421 LLDSETHPLGMAVKSS 437

RESULT 6
AA95975
ID AA95975 standard; Protein: 437 AA.
AC AA95975;
XX
DT 05-DEC-2000 (first entry)
XX
DE Mouse Sonic hedgehog Shh protein.
XX
KW Sonic hedgehog; Shh; mouse; agonist; antagonist;
KW lipid modulator; vacuole; cholesterol disorder; lipid disorder;
KW lipid metabolism; lipid storage; lipid transport; apolipoprotein;
KW triglyceride; hypercholesterolemia; abetalipoproteinemia;
KW hypobetalipoproteinemia; chylomicron retention; Anderson's disease;
KW fat absorption; atherosclerosis; obesity; weight loss;
KW vitamin A disorder; vitamin E disorder; anorexia; anorectic;
KW arteriosclerotic; gene therapy; diagnosis.
XX
OS Mus musculus.
XX
PN M020051628-A2.
XX
PD 08-SEP-2000.
XX
PF 03-MAR-2000; 2000MO-US05662.
XX
PR 03-MAR-1999; 99US-0122640.
PR 15-MAR-1999; 99US-0124446.
XX
PA (BIOJ) BIOGEN INC.
XX
PI Burkly L, Wang LC;
XX
DR WPL: 2000-611340/58.
DR N-PSDB; AAA50449.
XX
PT Use of lipid modulators (e.g. hedgehog agonists or antagonists) for
PT modulating lipid metabolism and storage, especially useful for treating
PT lipid metabolism or cholesterol disorders, e.g. obesity or
PT hypercholesterolemia -
XX
PS Disclosure: Page 113-115; 136pp; English.
XX
CC The present sequence of that of mouse Sonic hedgehog (Shh)
CC protein. The invention provides claimed methods for modulating
CC lipid metabolism, for modulating vacuole formation in intestinal
CC epithelial cells, for modulating the accumulation of fat in
CC intestinal epithelial cells, for treating a cholesterol disorder
CC and for treating a lipid metabolism disorder in an animal

Cc (especially a human) by administering a lipid modulator selected
Cc from a hedgehog antagonist or hedgehog agonist. In particular, the
Cc lipid metabolism disorder is a lipid storage disorder, a lipid
Cc transport disorder, an apolipoprotein disorder, a triglyceride
Cc disorder, e.g. a triglyceride metabolism disorder, a triglyceride
Cc transport disorder or a triglyceride storage disorder; a
Cc diet-induced hypercholesterolemia, hypercholesterolemia,
Cc abetalipoproteinemia, hypobetalipoproteinemia; a
Cc chylomicron-retention disorder, Anderson's disease, a fat
Cc absorption disorder, e.g. obesity or associated with weight loss,
Cc normotriglyceridemic abetalipoproteinemia, an apolipoprotein-B100
Cc deficiency, a fat soluble vitamin disorder, where the fat soluble
Cc vitamin is vitamin A or E, or atherosclerosis (all claimed). The
Cc hedgehog antagonist binds to the hedgehog receptor, but does not
Cc elicit a response. It is preferably a hedgehog mimetic, a modified
Cc hedgehog protein, e.g. an inactive hedgehog variant, or an
Cc anti-hedgehog homologue, especially a human, chimeric or humanised
Cc antibody. The methods are useful in preventing these disorders or
Cc protecting a subject from these disorders. The hedgehog antagonist
Cc and agonist are also useful in diagnosis and research associated
Cc with these disorders. The lipid modulators may also be used as a
Cc part of a gene therapy protocol to deliver polynucleotides encoding
Cc these lipid modulators.

Sequence 437 AA;

Query Match 100.0%; Score 2289; DB 21; Length 437;
Best Local Similarity 100.0%; Pred. No. 9,4e-226;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLLLRARCFVYLASSLLVCPGLACGPGRGFGKRRHPRKLTPLAYKQFIPNVAERTLGAS 60
Db 1 MLLLRARCFVYLASSLLVCPGLACGPGRGFGKRRHPRKLTPLAYKQFIPNVAERTLGAS 60
Qy 61 GRYEKKTRNSRERKELTPNNPDIFFDEENTGADRLMTQRCCKLNAISVNWMPG 120
Db 61 GRYEKKTRNSRERKELTPNNPDIFFDEENTGADRLMTQRCCKLNAISVNWMPG 120
Qy 121 VRLNRTGMDGHHSESLHYEGRAVDITTSDRRSKYGMARLAVEAGFDWYYESKA 180
Db 121 VRLNRTGMDGHHSESLHYEGRAVDITTSDRRSKYGMARLAVEAGFDWYYESKA 180
Qy 181 HHC5VKAENSVAAKSGCGCPGSAATVHLEGGCTKLVKDLRPGDRVLAADGGRLLYSDFL 240
Db 181 HHC5VKAENSVAAKSGCGCPGSAATVHLEGGCTKLVKDLRPGDRVLAADGGRLLYSDFL 240
Qy 241 TFLDRDEGAKKVFYIETLEPRERLLTAHLLFVAPHNDGSPPPGSPALFASRVPRQOR 300
Db 241 TFLDRDEGAKKVFYIETLEPRERLLTAHLLFVAPHNDGSPPPGSPALFASRVPRQOR 300
Qy 301 VYVAERGGDRRLPAAVHSVTLREBEAGAYAPLTAHGTILLINRYLASCYAVIEESHNAH 360
Db 301 VYVAERGGDRRLPAAVHSVTLREBEAGAYAPLTAHGTILLINRYLASCYAVIEESHNAH 360
Qy 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPRTAGIHMYSQLYHIGTW 420
Db 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPRTAGIHMYSQLYHIGTW 420
Qy 421 LLDSETHPLGMAVKSS 437
Db 421 LLDSETHPLGMAVKSS 437

RESULT 7
AA95284
ID AA95284 standard; Protein: 437 AA.
AC AA95284;
XX
DT 12-SEP-2000 (first entry)
XX
DE Mouse Sonic hedgehog Shh protein.
XX

KW Sonic hedgehog; Shh; mouse; excitotoxicity; Parkinson's disease;
 KW Huntington's disease; neuronal degeneration; neuroprotective;
 KW dopaminergic; GABAergic; substantia nigra; therapy.
 OS Mus musculus.
 PN W0200035948-A1.
 XX 22-JUN-2000.
 PD 03-DEC-1999; 99WO-US28721.
 PF 03-DEC-1999; 99WO-US28721.
 XX 03-DEC-1998; 98WO-US25676.
 PR 27-JAN-1999; 99US-0238243.
 PR 03-JUN-1999; 99US-0325602.
 XX (BIOJ) BIOGEN INC.
 PA (ONTO-) ONTOGENY INC.
 PI Galdes A, Mahanthappa N;
 DR WPI: 2000-431570/37.
 DR N-PSDB: AAA27879.
 XX Treating disorders involving exotoxicity, e.g. trauma, hypoglycemia,
 PT senile dementia and Korsakoff's disease, by using lipophilic modified
 PT hedgehog polypeptide -
 XX Disclosure: Page 154-155; 174pp: English.
 PS The present sequence of that of mouse Sonic hedgehog (Shh)
 CC protein. The invention relates to a method for promoting the
 CC survival and/or functional performance of neuronal cells,
 CC especially substantia nigra, dopaminergic or GABAergic neurons that
 CC are susceptible to exotoxicity, by contacting the cells, in vitro
 CC or in vivo, with a lipophilic (e.g. cholesterol) modified hedgehog
 CC polypeptide. The method is used to treat or prevent Parkinson's
 CC disease, Huntington's disease, domoic acid poisoning, spinal cord
 CC trauma, hypoglycemia, mechanical trauma to the nervous system,
 CC senile dementia, Korsakoff's disease, schizophrenia, AIDS dementia,
 CC multi-infarct dementia, mood disorders, depression, chemical
 CC toxicity, neuronal damage associated with uncontrolled seizures
 CC such as epileptic seizures, neuronal injury associated with HIV and
 CC AIDS, neurodegeneration associated with Down's syndrome,
 CC neuropathic pain syndrome, olivopontocerebral atrophy, amyotrophic
 CC lateral sclerosis, mitochondrial abnormalities, Alzheimer's disease,
 CC hepatic encephalopathy, Tourette's syndrome and drug addiction (all
 CC claimed). The lipophilic modified hedgehog polypeptide is also
 CC useful for promoting survival and/or functional performance of
 CC neuronal cells susceptible to exotoxicity.
 CC
 XX Sequence 437 AA:
 SQ
 Query Match 100.0%; Score 2289; DB 21; Length 437;
 Best Local Similarity 100.0%; Pred. No. 9,4e-226;
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 TFLDDEGAKVYVYETLEPERILLTAAHLLEVPAPNDGSPPPGSALEFASRVPCOR 300
 DB 241 TFLDDEGAKVYVYETLEPERILLTAAHLLEVPAPNDGSPPPGSALEFASRVPCOR 300
 QY 301 VYVAERGGDRRLPAAVHSTLREEGAVAPLTAGTTLINRLASCYAVIEHSNAH 360
 DB 301 VYVAERGGDRRLPAAVHSTLREEGAVAPLTAGTTLINRLASCYAVIEHSNAH 360
 QY 361 RAFAPEFLAHLAALAPARTDGGGGSIPAAQSTEARGAEPYAGIHMYSQLYHIGTW 420
 DB 361 RAFAPEFLAHLAALAPARTDGGGGSIPAAQSTEARGAEPYAGIHMYSQLYHIGTW 420
 QY 421 LLDSETHPLGMAYKSS 437
 DB 421 LLDSETHPLGMAYKSS 437
 RESULT 8
 ID AAY96246 standard; Protein: 437 AA.
 AC AAY96246;
 XX 11-SEP-2000 (first entry)
 DE Partial mouse Shh.
 XX Mouse; sonic hedgehog; Shh; neuromuscular disorder; neuropathy;
 KW Guillain-Barre syndrome; peripheral neuropathy; diabetes; alcoholism;
 KW chronic inflammatory demyelinating polyneuropathy; CIPD;
 KW gene therapy; infection; inflammation; hereditary neuropathy;
 KW Charcot-Marie-Tooth disease; vasculitis; lung cancer; tumour;
 KW multiple myeloma; nutritional imbalance; kidney disease;
 KW hypothyroid neuropathy; trauma; Refsum's disease; Abetalipoproteinemia;
 KW Fabry's disease; CMT; GBS; Dejerine-Sottas syndrome; leukodystrophy;
 KW amyotrophic lateral sclerosis; ALS; Miller-Fisher syndrome; acute neuropathy;
 KW hereditary sensory neuropathy type II; HSN II; B-cell lymphoma;
 KW Waldenstrom's Macroglobulinaemia; Chronic Lymphocytic Leukaemia;
 KW neuroprotective; cytoprotective; patched-mediated signal transduction.
 XX
 OS Mus sp.
 XX W0200027422-A2.
 PN 18-MAY-2000.
 XX 08-NOV-1999; 99WO-US26334.
 PF 06-NOV-1998; 98US-0187387.
 PR (BIOJ) BIOGEN INC.
 PA (ONTO-) ONTOGENY INC.
 PI Galdes A, Mahanthappa N;
 DR WPI: 2000-387341/33.
 DR N-PSDB: AAA30277.
 XX Novel method of preventing deterioration of peripheral nerves, useful
 PT for treating or preventing neuropathy, e.g. where associated with
 PT diabetes or viral infection, by administering hedgehog or patched agent
 XX
 PS Claim 7: Page 131-132; 152pp: English.
 XX The present sequence is the partial mouse sonic hedgehog protein, Shh.
 CC This sequence inhibits expression of the patched gene which has been
 CC implicated in neuromuscular disorders (neuropathies). This sequence may
 CC therefore be used for treating neuromuscular disorders i.e. preventing
 CC degeneration in function of motor or sensory nerves and protecting
 CC peripheral nerve cells under conditions that normally cause neuropathy.
 CC A variety of neuromuscular disorders may be treated: Guillain-Barre

CC syndrome, GBS, peripheral neuropathy; diabetic neuropathy;
CC alcohol-induced neuropathy; chronic inflammatory demyelinating
CC polyneuropathy, CIDP; infection-induced neuropathy, including HIV
CC infection; inflammation-induced neuropathy; hereditary neuropathy e.g.
CC Charcot-Marie-Tooth disease (CMT), Familial Amyloidotic neuropathy,
CC Refsum's disease, Abetalipoproteinemia, Tangier disease, Krabbe's
CC disease, Metachromatic leukodystrophy, Fabry's disease, Dejerine-Sottas
CC syndrome, Hereditary sensory neuropathy type II (HSN II) and Amyotrophic
CC lateral sclerosis (ALS); acute neuropathy e.g Miller-Fisher syndrome;
CC neuropathy caused by vasculitis; neuropathy associated with tumours e.g.
CC lung cancer, multiple myeloma, B-cell lymphoma, Waldenstrom's
CC Macroglobulinaemia, Chronic lymphocytic Leukaemia, neuropathy associated
CC with: amyloidosis, nutritional imbalance, kidney disease, trauma; and
CC hypothyroid neuropathy. The coding sequence may be used in gene therapy
of the above disorders.

SO Sequence 437 AA:

Query Match 100.0%; Score 2289; DB 21; Length 437;
Best Local Similarity 100.0%; Pred. No. 9,4e-226;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLLLLARCFVLIASSLLVCPGLACGPGRGFGKRRHKLTPLAYKOFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIASSLLVCPGLACGPGRGFGKRRHKLTPLAYKOFIPNVAEKTIGAS 60
OY 61 GRYEGKITRNSERFKELTPYNYNDIIFKDEENTGADRLMTORCKDKMLALATISVMNQMPG 120
DB 61 GRYEGKITRNSERFKELTPYNYNDIIFKDEENTGADRLMTORCKDKMLALATISVMNQMPG 120
OY 121 VRLRVTEGMDGDHSEESLHYEGRAVDITTSDBRSKYGMLARLAVEAGFDWVYESKA 180
DB 121 VRLRVTEGMDGDHSEESLHYEGRAVDITTSDBRSKYGMLARLAVEAGFDWVYESKA 180
OY 181 HHCSSVAENSVAKSGCGCPGSAVYHLBOGRTKLVYDLRPGDRVLAADOGRLYSDFL 240
DB 181 HHCSSVAENSVAKSGCGCPGSAVYHLBOGRTKLVYDLRPGDRVLAADOGRLYSDFL 240
OY 241 TFLDRDEGAKKVFYITLTPRERLLTAHLLEFVAPHNDSGPTPGSALFASVRGQR 300
DB 241 TFLDRDEGAKKVFYITLTPRERLLTAHLLEFVAPHNDSGPTPGSALFASVRGQR 300
OY 301 VYVVAERGGRRLLPAVHVSVTLRBEAGAYAPLTAHGTLLINRVLASCAVIEHSMW 360
DB 301 VYVVAERGGRRLLPAVHVSVTLRBEAGAYAPLTAHGTLLINRVLASCAVIEHSMW 360
OY 361 RAFAPRRLAHLAALAPARTDGGGGSIPAAOSATFARCAEPAGIHWYSOLLYHIGTW 420
DB 361 RAFAPRRLAHLAALAPARTDGGGGSIPAAOSATFARCAEPAGIHWYSOLLYHIGTW 420
OY 421 LLDSETHMPLGMAVKSS 437
DB 421 LLDSETHMPLGMAVKSS 437

RESULT 9
AA70679
ID AAY70679 standard; Protein; 437 AA.
XX

AC AAY70679;

DT 18-JUL-2000 (first entry)

XX Mouse Sonic hedgehog (Shh) protein.

KM Sonic hedgehog; Shh; mouse; growth modulator; therapeutic agent;
KM lung; hedgehog; hh; patched; ptc; fibroblast growth factor; fgf-10;
KM antiproliferative; anticancer; vulnerability; antineoplastic; hypotensive;
KM anti-inflammatory; antiasthmatic; antiarthritic; tuberculostatic; asthma;
KM antimicrobial; antifungal; treatment; prevention; lung diseases; cancer;
KM cystic fibrosis; bronchitis; emphysema; respiratory distress syndrome;
KM tuberculosi; wound healing; lung transplantation.

XX

OS Mus sp.
PN WO200015246-A2.
XX 23-MAR-2000.
PD 10-SEP-1999; 99MO-US20500.
XX 11-SEP-1998; 98US-0099952.
PR (HARD) HARVARD COLLEGE.
XX Pepticell C, Lewis P, McMahon AP;
XX WPI: 2000-271252/23.
DR N-PSDB; AA252260.
XX

PT Modulation of lung tissue or cell growth rate used for treating or
PT preventing damage to lung tissue comprises ectopically contacting
PT tissue with hedgehog therapeutic, patched therapeutic or fibroblast
PT growth factor-10

PS Claim 14; page 119-121; 143pp; English.

CC The patent discloses a method for modulating the growth state of
CC epithelial or mesenchymal cells of the lung, by ectopically contacting
CC the tissue with a therapeutic agent, that can effectively alter the rate
CC of proliferation of cells. This agent can be selected from hedgehog (hh),
CC patched (ptc) or fibroblast growth factor (fgf)-10 therapeutics. It
CC involves a direct or indirect antagonism of patched-mediated regulation
CC of gene expression. This method is useful for the treatment or prevention
CC of lung diseases, like cancer, cystic fibrosis, bronchopneumoconiosis,
CC bronchitis, bronchospasm, sarcoidosis, silicosis, eosinophilic granuloma,
CC ankylosing spondylitis, emphysema, tuberculosis, respiratory distress
CC syndrome, allergic rhinitis, asthma, pulmonary fibrosis and primary
CC pulmonary hypertension. It is also used to control wound healing or other
CC reformation processes in the lung and augment lung transplantation. The
CC present sequence is the mouse sonic hedgehog (Shh) protein, essential
CC for development of the respiratory system. Hedgehog polypeptides can be
CC used to control the formation and/or maintenance of the lung tissue.

SO Sequence 437 AA:

Query Match 100.0%; Score 2289; DB 21; Length 437;
Best Local Similarity 100.0%; Pred. No. 9,4e-226;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLLLLARCFVLIASSLLVCPGLACGPGRGFGKRRHKLTPLAYKOFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIASSLLVCPGLACGPGRGFGKRRHKLTPLAYKOFIPNVAEKTIGAS 60
OY 61 GRYEGKITRNSERFKELTPYNYNDIIFKDEENTGADRLMTORCKDKMLALATISVMNQMPG 120
DB 61 GRYEGKITRNSERFKELTPYNYNDIIFKDEENTGADRLMTORCKDKMLALATISVMNQMPG 120
OY 121 VRLRVTEGMDGDHSEESLHYEGRAVDITTSDBRSKYGMLARLAVEAGFDWVYESKA 180
DB 121 VRLRVTEGMDGDHSEESLHYEGRAVDITTSDBRSKYGMLARLAVEAGFDWVYESKA 180
OY 181 HHCSSVAENSVAKSGCGCPGSAVYHLBOGRTKLVYDLRPGDRVLAADOGRLYSDFL 240
DB 181 HHCSSVAENSVAKSGCGCPGSAVYHLBOGRTKLVYDLRPGDRVLAADOGRLYSDFL 240
OY 241 TFLDRDEGAKKVFYITLTPRERLLTAHLLEFVAPHNDSGPTPGSALFASVRGQR 300
DB 241 TFLDRDEGAKKVFYITLTPRERLLTAHLLEFVAPHNDSGPTPGSALFASVRGQR 300
OY 301 VYVVAERGGRRLLPAVHVSVTLRBEAGAYAPLTAHGTLLINRVLASCAVIEHSMW 360
DB 301 VYVVAERGGRRLLPAVHVSVTLRBEAGAYAPLTAHGTLLINRVLASCAVIEHSMW 360
OY 361 RAFAPRRLAHLAALAPARTDGGGGSIPAAOSATFARCAEPAGIHWYSOLLYHIGTW 420
DB 361 RAFAPRRLAHLAALAPARTDGGGGSIPAAOSATFARCAEPAGIHWYSOLLYHIGTW 420

121 VRLRVTEGWDEGHHSSESLHYEGRAVDITTSDBRDSKYGMABIAVENCEDTIVVDR 100

particularly chondrocytes, or testicular germ line cells. The present

sequence represents a mouse Shh polypeptide.

SQ Sequence 437 AA: 100.0%; Score 2289; DB 22; Length 437;
 Query Match Best Local Similarity 100.0%; Pred. No. 9.4e-226;
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLLIARCFVLIASSLLVCPGLACGPGRGGRKRRHKRLTPLAYKQFIPNVAEKTIGAS 60
 DB 1 MLLIARCFVLIASSLLVCPGLACGPGRGGRKRRHKRLTPLAYKQFIPNVAEKTIGAS 60
 QY 61 GRYEGKITRNSERKELTPNPNPDIIFKDEENTGADRLMTORCKDKLNALISVNMQWPG 120
 DB 61 GRYEGKITRNSERKELTPNPNPDIIFKDEENTGADRLMTORCKDKLNALISVNMQWPG 120
 QY 121 VRLRTGEGWDEGHHSEESLHYEGRAVDITTSDBRSKYGMRLARLAVEAGFDWYYESKA 180
 DB 121 VRLRTGEGWDEGHHSEESLHYEGRAVDITTSDBRSKYGMRLARLAVEAGFDWYYESKA 180
 QY 181 HIRCSVKAENSVAAKSGGCFPGSATVHLEOGGTLYKDLRPGDRVLADDDGRLYSDFL 240
 DB 181 HIRCSVKAENSVAAKSGGCFPGSATVHLEOGGTLYKDLRPGDRVLADDDGRLYSDFL 240
 QY 241 TFLDRDEGAKKVFYIETLEPERERLLTAAHLLEFVAPHNDSGPTGPSALFASRVPRGQR 300
 DB 241 TFLDRDEGAKKVFYIETLEPERERLLTAAHLLEFVAPHNDSGPTGPSALFASRVPRGQR 300
 QY 301 VYVVAERGGDRRLPAVAHSVTLREEBAGAYAPLTAHGTILINRVLASCYAVIEESHMAH 360
 DB 301 VYVVAERGGDRRLPAVAHSVTLREEBAGAYAPLTAHGTILINRVLASCYAVIEESHMAH 360
 QY 361 RAFAFRLAHALLAALAPARTDGGGGSTIPAQOSATEARGAEPAGIHWYSOLLYHIGTW 420
 DB 361 RAFAFRLAHALLAALAPARTDGGGGSTIPAQOSATEARGAEPAGIHWYSOLLYHIGTW 420
 QY 421 LLDSETHPLGMAVKSS 437
 DB 421 LLDSETHPLGMAVKSS 437
 RESULT 12
 AAB84672
 ID AAB84672 standard; Protein; 437 AA.
 AC AAB84672;
 XX
 DT 17-SEP-2001 (first entry)
 XX
 DE Amino acid sequence of a mouse hedgehog (Shh) polypeptide.
 XX
 KW Shh gene; hedgehog gene; T lymphocyte; patched gene; infection;
 KW diabetes; nutritional deficiency; graft rejection; hyperacute response;
 KW cornea transplant; autoimmune disorder; multiple sclerosis; psoriasis;
 KW atopic dermatitis; inflammatory disease; proliferative disease;
 KW hyperproliferative disease; eczematous dermatitis; urticaria;
 KW vasculitis; scleroderma.
 XX
 OS Mus sp.
 XX
 PN MO20010438-A2.
 XX
 PD 07-JUN-2001.
 XX
 PF 30-NOV-2000; 2000MO-US32590.
 XX
 PR 30-NOV-1999; 99US-0168112.
 XX
 PA (CUR1-) CURIS INC.
 XX
 PI Crompton T;
 XX
 DR WPI: 2001-44184/47.
 DR N-PSDB; AAB84672.
 XX

PT Modulating immune function comprises administration of a hedgehog or
 PT ptc agent, for treating e.g. diabetes, eczematous dermatitis, urticaria
 PT or vasculitis -
 XX
 PS Claim 4; Page 81-82; 105pp; English.
 CC The present sequence represents a hedgehog (Shh) polypeptide. Hedgehog
 CC gene products and signal transduction pathways involving hedgehog are
 CC involved in the maturation of T lymphocytes. The specification describes
 CC a method for modulating immune function, by administration of a hedgehog
 CC or patched (ptc) polypeptide, agonists or antagonists. The method is
 CC used to treat disorders affecting the regulation of lymphocytes,
 CC particularly maturation and/or activation of T lymphocytes. It is used
 CC to treat bacterial or viral infection, diabetes, nutritional
 CC deficiencies, graft rejection or other hyperacute response such as
 CC kidney, heart, lung, bone marrow spleen skin or cornea transplant or
 CC autoimmune disorders such as multiple sclerosis, psoriasis or atopic
 CC dermatitis. The method is used to treat inflammatory, proliferative and
 CC hyperproliferative diseases, as well as cutaneous manifestations of
 CC immunological disorders such as eczematous dermatitis, urticaria,
 CC vasculitis and scleroderma.
 CC
 SQ Sequence 437 AA: 100.0%; Score 2289; DB 22; Length 437;
 Query Match Best Local Similarity 100.0%; Pred. No. 9.4e-226;
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLLIARCFVLIASSLLVCPGLACGPGRGGRKRRHKRLTPLAYKQFIPNVAEKTIGAS 60
 DB 1 MLLIARCFVLIASSLLVCPGLACGPGRGGRKRRHKRLTPLAYKQFIPNVAEKTIGAS 60
 QY 61 GRYEGKITRNSERKELTPNPNPDIIFKDEENTGADRLMTORCKDKLNALISVNMQWPG 120
 DB 61 GRYEGKITRNSERKELTPNPNPDIIFKDEENTGADRLMTORCKDKLNALISVNMQWPG 120
 QY 121 VRLRTGEGWDEGHHSEESLHYEGRAVDITTSDBRSKYGMRLARLAVEAGFDWYYESKA 180
 DB 121 VRLRTGEGWDEGHHSEESLHYEGRAVDITTSDBRSKYGMRLARLAVEAGFDWYYESKA 180
 QY 181 HIRCSVKAENSVAAKSGGCFPGSATVHLEOGGTLYKDLRPGDRVLADDDGRLYSDFL 240
 DB 181 HIRCSVKAENSVAAKSGGCFPGSATVHLEOGGTLYKDLRPGDRVLADDDGRLYSDFL 240
 QY 241 TFLDRDEGAKKVFYIETLEPERERLLTAAHLLEFVAPHNDSGPTGPSALFASRVPRGQR 300
 DB 241 TFLDRDEGAKKVFYIETLEPERERLLTAAHLLEFVAPHNDSGPTGPSALFASRVPRGQR 300
 QY 301 VYVVAERGGDRRLPAVAHSVTLREEBAGAYAPLTAHGTILINRVLASCYAVIEESHMAH 360
 DB 301 VYVVAERGGDRRLPAVAHSVTLREEBAGAYAPLTAHGTILINRVLASCYAVIEESHMAH 360
 QY 361 RAFAFRLAHALLAALAPARTDGGGGSTIPAQOSATEARGAEPAGIHWYSOLLYHIGTW 420
 DB 361 RAFAFRLAHALLAALAPARTDGGGGSTIPAQOSATEARGAEPAGIHWYSOLLYHIGTW 420
 QY 421 LLDSETHPLGMAVKSS 437
 DB 421 LLDSETHPLGMAVKSS 437
 RESULT 13
 AAE04685
 ID AAE04685 standard; Protein; 437 AA.
 AC AAE04685;
 XX
 DT 04-SEP-2001 (first entry)
 XX
 DE Mouse sonic hedgehog (Shh) protein.
 XX
 KW Mouse; hedgehog protein; noctropic; neuroprotective; anticonvulsant;
 KW cytosstatic; therapy; Alzheimer's disease; Parkinson's disease; injury;
 KW

KW Huntington's chorea; amyotrophic lateral sclerosis; multiple sclerosis;
 KW nervous system aging; neurodegenerative disease; immunological disease;
 KW malignant glioma; medulloblastoma; neuroectodermal tumour; cancer;
 XX extracellular signalling protein.
 OS Mus sp.
 PN WO200134654-A1.
 XX 17-MAY-2001.
 PD 02-NOV-2000; 2000WO-US30405.
 XX 05-NOV-1999; 99US-0164025.
 PR (BioJ) BIOGEN INC.
 XX Strauch K;
 PI WPI: 2001-329075/34.
 DR N-PSDB; AAD09032.
 XX
 PT Novel isolated hedgehog fusion polypeptide useful for treating
 PT neurological conditions such as Alzheimer's disease, Parkinson's
 PT disease, Huntington's chorea, amyotrophic lateral sclerosis, and
 XX multiple sclerosis.
 PS Disclosure; Page 115-116; 178pp; English.
 CC The present invention relates to hedgehog fusion proteins. Hedgehog
 CC proteins are a family of extracellular signalling proteins that regulate
 CC various aspects of embryonic development both in vertebrates and in
 CC invertebrates. Hedgehog fusion protein is useful for the prophylaxis or
 CC treatment of any condition or disease state for which a hedgehog or
 CC patched protein constituent is efficacious and in the diagnosis of
 CC constituents or conditions of disease states and in the diagnosis of
 CC specimens and for diagnostic purposes in non-physiological systems or
 CC Hedgehog fusion protein is useful for treating neurological conditions
 CC due to injury, aging of nervous system, including Alzheimer's disease,
 CC chronic neurodegenerative diseases of the nervous system, including
 CC Parkinson's disease, Huntington's chorea, amyotrophic lateral sclerosis,
 CC and chronic immunological diseases of the nervous system including
 CC sclerosis and malignant gliomas, medulloblastomas, neuroectodermal
 CC tumours and to specifically target medical therapies against cancers and
 CC tumours which express the receptor for the protein. The present sequence
 CC is mouse sonic hedgehog (Shh) protein.
 CC
 CC Sequence 437 AA:
 SQ
 Query Match 100.0%; Score 2289; DB 22; Length 437;
 Best Local Similarity 100.0%; Pred. No. 9,4e-226;
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLLILARCFVILASSLLVCPGLACGPGRGFGKRRHPRKLLPLAVYKQFIPNVAEKTIGAS 60
 DB 1 MLLILARCFVILASSLLVCPGLACGPGRGFGKRRHPRKLLPLAVYKQFIPNVAEKTIGAS 60
 QY 61 GRYEGKTRNRSERKELTPNPNPDIIFKDEMGADRLMTORCKDKMALAISVMQWPG 120
 DB 61 GRYEGKTRNRSERKELTPNPNPDIIFKDEMGADRLMTORCKDKMALAISVMQWPG 120
 QY 121 VRLRVTEGMDHSEESLHYEGRAVDITTSRDSKSKGMALARLAVEGFDMVYVESKA 180
 DB 121 VRLRVTEGMDHSEESLHYEGRAVDITTSRDSKSKGMALARLAVEGFDMVYVESKA 180
 QY 181 HHCSTKANSVAANSKGCGPGSATVHLEGGTKLVKDLRPGDRLVLAADQGRLLYSDPL 240
 DB 181 HHCSTKANSVAANSKGCGPGSATVHLEGGTKLVKDLRPGDRLVLAADQGRLLYSDPL 240
 QY 241 TFLDDEGAKKVFYIETLEPERLLTAHLLEFVAPHNDSPPTGSPALFASRVPRGQR 300
 DB 241 TFLDDEGAKKVFYIETLEPERLLTAHLLEFVAPHNDSPPTGSPALFASRVPRGQR 300

QY 301 VYVVAERGDRRLPLPAVHSVTLREERAGAYAPLTAHGTILINRVLASCYAVEESMAH 360
 DB 301 VYVVAERGDRRLPLPAVHSVTLREERAGAYAPLTAHGTILINRVLASCYAVEESMAH 360
 QY 361 RAFAFRLAHALLAALAPARTDGGGGSTPAQASATREANGAEPPTAGIHWSQLLYHIGTW 420
 DB 361 RAFAFRLAHALLAALAPARTDGGGGSTPAQASATREANGAEPPTAGIHWSQLLYHIGTW 420
 QY 421 LLDSETHMPLGMAVKS 437
 DB 421 LLDSETHMPLGMAVKS 437
 RESULT 14
 AAE05375
 ID AAE05375 standard; Protein: 437 AA.
 XX
 AC AAE05375;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Mouse Sonic hedgehog protein.
 XX
 KW Mouse; Sonic hedgehog; Shh; morphogenic signal; neuron; chromosome 5;
 KW embryonic patterning; cell culture; cell differentiation; ischaemia;
 KW cell proliferative disorder; intracerebral grafting; Huntington's chorea;
 KW amyotrophic lateral sclerosis; ALS; multiple sclerosis.
 OS Mus sp.
 XX
 PH Key Location/Qualifiers
 FT Peptide 1..24
 FT Protein /label=Signal_peptide
 FT /label=Signal_peptide
 XX /label=Mature_Shh_protein
 XX US6261786-B1.
 XX 17-JUL-2001.
 PD 02-JUL-1996; 96US-0674509.
 PF 30-DEC-1993; 93US-0176427.
 PR 14-DEC-1994; 94US-0356060.
 PR 04-MAY-1995; 95US-0435093.
 PR 05-JUN-1995; 95US-0460900.
 PR 05-JUN-1995; 95US-0462386.
 XX
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
 XX (HARD) HARVARD COLLEGE.
 PI Mariago V, Tablin CJ, Ingham PW, McMahon AP;
 XX WPI: 2001-440859/47.
 DR N-PSDB; AAD10149.
 XX
 PT Screening compounds that potentiate or inhibit binding of hedgehog
 PT polypeptide to naturally occurring patched receptor, comprises
 PT contacting polypeptide with receptor and test compound, and detecting
 PT change in binding.
 PS Claim 2; Fig 5A; 127pp; English.
 XX
 CC The present invention relates to assay for screening compounds that
 CC potentiate or inhibit binding of hedgehog polypeptide to naturally
 CC occurring patched receptor. The hedgehog proteins comprise morphogenic
 CC signals produced by embryonic patterning centres, and are involved in
 CC formation and maintenance of ordered spatial arrangements of
 CC differentiated tissues in vertebrates, both adult and embryonic. The
 CC proteins can be used to generate and/or maintain an array of different
 CC vertebrate tissues both in vitro and in vivo. The invention also relates
 CC to a method for modulating growth, differentiation or survival of a

CC mammalian cell (e.g. neuron, testicular cell) responsive to hedgehog
 CC induction. Hedgehog agonists and antagonists can be used in cell culture
 CC techniques to enhance survival and maintenance of neurons and various
 CC vertebrate organogenic pathways. The hedgehog gene is useful in
 CC determining whether a patient is at the risk of disorder characterised by
 CC unwanted cell proliferation or aberrant control of differentiation. The
 CC hedgehog proteins or mimetics can be used to induce fetal neurons
 CC especially neuronal stem cells in intracerebral grafting. The protein
 CC or its mimetic can be used in the treatment of neurological conditions
 CC e.g. injury to nervous system, ischaemia resulting from stroke,
 CC Alzheimer's disease, Parkinson's disease, Huntington's chorea,
 CC amyotrophic lateral sclerosis (ALS) and multiple sclerosis. The present
 CC sequence is mouse Sonic hedgehog (Shh) protein. The Shh gene is
 CC mapped to the proximal region of mouse chromosome 5.

XX Sequence 437 AA:

Query Match 100.0%: Score 2289; DB 22; Length 437;

Best Local Similarity 100.0%: Pred. No. 9.4e-226; Mismatches 0; Indels 0; Gaps 0;

Matches 437: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLAARCFVLVIASSLVCPLACGPGRGFGRRHPPKLTPLAYKQFIIPNVAEKTIGAS 60
 DB 1 MLLLAARCFVLVIASSLVCPLACGPGRGFGRRHPPKLTPLAYKQFIIPNVAEKTIGAS 60

QY 61 GRYEGKITRNSERFKELTNPYNPDIIFKDEENTGADRLMTORCKDKLNLALISVNMQMG 120
 DB 61 GRYEGKITRNSERFKELTNPYNPDIIFKDEENTGADRLMTORCKDKLNLALISVNMQMG 120

QY 121 VRLRTYEGWDEGHHSEESLHYEGRAVDITTSDRDSKTKGMLARLAVEGFDMVYTESKA 180
 DB 121 VRLRTYEGWDEGHHSEESLHYEGRAVDITTSDRDSKTKGMLARLAVEGFDMVYTESKA 180

QY 181 HTHCSYKAENSYAASGGCGFPGSATVHLBOGGTKLVKDLRPGDRVLAADOGRLYSDFL 240
 DB 181 HTHCSYKAENSYAASGGCGFPGSATVHLBOGGTKLVKDLRPGDRVLAADOGRLYSDFL 240

QY 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLFVAAPHNDSGPTPGPSALFASRVAPGQR 300
 DB 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLFVAAPHNDSGPTPGPSALFASRVAPGQR 300

QY 301 VYVVAERGGDRRLPRAAVHSVTLREBEAGAYAPLTAHGTILINRYLASCYAVIEESHMAH 360
 DB 301 VYVVAERGGDRRLPRAAVHSVTLREBEAGAYAPLTAHGTILINRYLASCYAVIEESHMAH 360

QY 361 RAFAPERLAHLAALAPARTDGGGGSIPAAQSATGARGAEPAGIHMYSQLYHIGTW 420
 DB 361 RAFAPERLAHLAALAPARTDGGGGSIPAAQSATGARGAEPAGIHMYSQLYHIGTW 420

QY 421 LLDSETHMPLGMAVKSS 437
 DB 421 LLDSETHMPLGMAVKSS 437

RESULT 15
 AAB31220
 ID AAB31220 standard; Protein; 437 AA.
 AC AAB31220;
 XX 20-APR-2001 (first entry)
 DE Amino acid sequence of mouse sonic hedgehog protein (Shh).
 KW Hedgehog related-protein; sonic hedgehog protein; Shh; ischemia; stroke;
 KW desert hedgehog protein; Dhh; Indian hedgehog protein; Ihh; neuron;
 KW neurological condition; nervous system injury; tumour-induced injury;
 KW aging; Alzheimer's disease; chronic neurodegenerative disease;
 KW Parkinson's disease; Huntington's chorea; amyotrophic lateral sclerosis;
 KW spinocerebellar degeneration; chronic immunological disease;
 XX multiple sclerosis.
 OS Mus sp.

XX US6165747-A.
 PN 26-DEC-2000.
 PD 05-JUN-1995; 95US-0460900.
 XX 30-DEC-1993; 93US-0176427.
 PR 14-DEC-1994; 94US-0356060.
 PR 04-MAY-1995; 95US-0435093.
 XX (HARD) HARVARD COLLEGE.
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
 XX Ingham PW, McMahon AP, Tablin CJ, Marti-gorostiza E, Bumcrot DA;
 PI MPI: 2001-079847/09.
 DR N-PSDB: MAC87077.
 XX Polynucleotides encoding hedgehog proteins, useful for treating
 PT diseases of nervous system such as Alzheimer's disease, Parkinson's
 PT disease, Huntington's chorea, amyotrophic lateral sclerosis, multiple
 PT sclerosis -
 PS Claim 10; Columns 143-146; 119pp; English.

The present sequence represents a hedgehog related-protein. The
 specification describes a sonic hedgehog protein (Shh), a desert
 hedgehog protein (Dhh), and an Indian hedgehog protein (Ihh). The
 hedgehog polynucleotides are useful in diagnostic, in antisense
 therapy, and in therapeutic assays for detecting and treating disorders
 involving, e.g., aberrant expression of vertebrate hedgehog homologue.
 CC Hedgehog polypeptides are useful therapeutically to enhance survival
 CC of neurons and other neuron cells and in treating neurological
 CC conditions deriving from acute, subacute, or chronic injury to the
 CC nervous system, including traumatic injury, chemical injury, vascular
 CC injury and deficits (such as the ischemia resulting from stroke),
 CC together with infectious/inflammatory and induced-induced injury, aging
 CC of the nervous system including Alzheimer's disease, chronic
 CC neurodegenerative diseases of the nervous system, including Parkinson's
 CC disease, Huntington's chorea, amyotrophic lateral sclerosis,
 CC spinocerebellar degenerations, and chronic immunological diseases of
 CC the nervous system or affecting the nervous system, including multiple
 CC sclerosis.

XX Sequence 437 AA:

Query Match 100.0%: Score 2289; DB 22; Length 437;

Best Local Similarity 100.0%: Pred. No. 9.4e-226; Mismatches 0; Indels 0; Gaps 0;

Matches 437: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLAARCFVLVIASSLVCPLACGPGRGFGRRHPPKLTPLAYKQFIIPNVAEKTIGAS 60
 DB 1 MLLLAARCFVLVIASSLVCPLACGPGRGFGRRHPPKLTPLAYKQFIIPNVAEKTIGAS 60

QY 61 GRYEGKITRNSERFKELTNPYNPDIIFKDEENTGADRLMTORCKDKLNLALISVNMQMG 120
 DB 61 GRYEGKITRNSERFKELTNPYNPDIIFKDEENTGADRLMTORCKDKLNLALISVNMQMG 120

QY 121 VRLRTYEGWDEGHHSEESLHYEGRAVDITTSDRDSKTKGMLARLAVEGFDMVYTESKA 180
 DB 121 VRLRTYEGWDEGHHSEESLHYEGRAVDITTSDRDSKTKGMLARLAVEGFDMVYTESKA 180

QY 181 HTHCSYKAENSYAASGGCGFPGSATVHLBOGGTKLVKDLRPGDRVLAADOGRLYSDFL 240
 DB 181 HTHCSYKAENSYAASGGCGFPGSATVHLBOGGTKLVKDLRPGDRVLAADOGRLYSDFL 240

QY 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLFVAAPHNDSGPTPGPSALFASRVAPGQR 300
 DB 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLFVAAPHNDSGPTPGPSALFASRVAPGQR 300

QY 301 VYVVAERGGDRRLPRAAVHSVTLREBEAGAYAPLTAHGTILINRYLASCYAVIEESHMAH 360
 DB 301 VYVVAERGGDRRLPRAAVHSVTLREBEAGAYAPLTAHGTILINRYLASCYAVIEESHMAH 360

Db 301 VYVAERCGDRRLPAVHSTVLRREEAGAVPLTAHCTILINRVIASCYAVIEHSMH 360
Oy 361 RAFAFERLAHALALAPARTDGGGGSIPMAOSATEARGAEPAGIHWSOLYHGTW 420
Db 361 RAFAFERLAHALALAPARTDGGGGSIPMAOSATEARGAEPAGIHWSOLYHGTW 420
Oy 421 LLDSETHMPLGMAVKSS 437
Db 421 LLDSETHMPLGMAVKSS 437

Search completed: February 20, 2003, 10:10:18
Job time : 32.7046 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:08:38 ; Search time 11.4398 Seconds
(without alignments)
1123.956 Million cell updates/sec

Title: US-09-827-110A-13
Perfect score: 2289
Sequence: 1 MILLARCFLVILASSLLVC.....GTWLDSETHMPLGMAYKSS 437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2289	100.0	437	1 US-08-176-427B-8	Sequence 8, Appl1
2	2289	100.0	437	2 US-08-356-060A-11	Sequence 11, Appl1
3	2289	100.0	437	4 US-08-460-900C-11	Sequence 11, Appl1
4	2289	100.0	437	4 US-08-674-509B-11	Sequence 11, Appl1
5	2289	100.0	437	4 US-08-954-698-11	Sequence 11, Appl1
6	2289	100.0	437	4 US-08-957-874-11	Sequence 11, Appl1
7	2289	100.0	437	4 US-09-325-256-20	Sequence 20, Appl1
8	2289	100.0	437	3 US-08-946-329A-20	Sequence 20, Appl1
9	2289	99.9	437	4 US-08-567-357A-20	Sequence 20, Appl1
10	2289	99.9	437	4 US-08-729-743A-20	Sequence 20, Appl1
11	2289	99.9	437	4 US-09-057-860A-6	Sequence 6, Appl1
12	2289	99.9	437	4 US-08-349-498-20	Sequence 20, Appl1
13	2289	99.9	437	4 US-08-293-505-14	Sequence 14, Appl1
14	2289	99.9	437	5 PCT-US95-15463-20	Sequence 20, Appl1
15	2289	99.9	437	5 PCT-US95-15923-20	Sequence 20, Appl1
16	2289	99.9	437	4 US-08-757-230A-2	Sequence 2, Appl1
17	2289	96.1	437	4 US-08-757-230A-9	Sequence 9, Appl1
18	2289	96.1	437	5 PCT-US95-02315-2	Sequence 4, Appl1
19	2289	90.2	462	1 US-08-748-591-4	Sequence 9, Appl1
20	2289	90.2	462	1 US-08-356-060A-13	Sequence 13, Appl1
21	2289	90.2	475	2 US-08-460-900C-13	Sequence 13, Appl1
22	2289	90.2	475	4 US-08-674-509B-13	Sequence 13, Appl1
23	2289	90.2	475	4 US-08-954-698-13	Sequence 13, Appl1
24	2289	90.2	475	4 US-08-957-874-13	Sequence 13, Appl1
25	2289	90.2	475	4 US-09-325-256-22	Sequence 22, Appl1
26	2289	90.2	475	4 US-08-349-498-18	Sequence 18, Appl1
27	2289	78.1	425	1 US-08-176-427B-2	Sequence 2, Appl1

28	1788	78.1	425	2 US-08-356-060A-8	Sequence 8, Appl1
29	1788	78.1	425	4 US-08-460-900C-8	Sequence 8, Appl1
30	1788	78.1	425	4 US-08-674-509B-8	Sequence 8, Appl1
31	1788	78.1	425	4 US-08-954-698-8	Sequence 8, Appl1
32	1788	78.1	425	4 US-08-957-874-8	Sequence 8, Appl1
33	1788	78.1	425	4 US-09-325-256-17	Sequence 17, Appl1
34	1679	73.4	425	3 US-08-567-357A-19	Sequence 19, Appl1
35	1679	73.4	425	4 US-08-729-743A-19	Sequence 19, Appl1
36	1679	73.4	425	4 US-08-349-498-19	Sequence 19, Appl1
37	1679	73.4	425	4 US-08-954-698-19	Sequence 19, Appl1
38	1679	73.4	425	5 PCT-US95-15463-19	Sequence 19, Appl1
39	1679	73.4	425	5 PCT-US95-15923-19	Sequence 19, Appl1
40	1517	66.3	418	3 US-08-946-329A-18	Sequence 18, Appl1
41	1517	66.3	418	4 US-08-567-357A-18	Sequence 18, Appl1
42	1517	66.3	418	4 US-08-729-743A-18	Sequence 18, Appl1
43	1517	66.3	418	4 US-08-757-230A-7	Sequence 7, Appl1
44	1517	66.3	418	4 US-08-349-498-18	Sequence 18, Appl1
45	1517	66.3	418	5 PCT-US95-15463-18	Sequence 18, Appl1

ALIGNMENTS

RESULT 1
US-08-176-427B-8
; Sequence 8, Application US/08176427B
; Patent No. 5789543
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESSEE: LAHYE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-176-427B-8
Query Match 100.0% Score 2289; DB 1; Length 437;
Best local Similarity 100.0%; Pred. No. 2.7e-249; Indels 0; Gaps 0;
Matches 437; Conservative 0; Mismatches 0;

QY 1 MILLARCFLVILASSLLVCPGLACGPGRGFRHKKLTPLAYKQFIIPNVAEKLGLAS 60
DB 1 MILLARCFLVILASSLLVCPGLACGPGRGFRHKKLTPLAYKQFIIPNVAEKLGLAS 60

QY 61 GYEGKITRNSERFELTPNYNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMWPG 120
 DB 61 GYEGKITRNSERFELTPNYNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMWPG 120
 QY 121 VRLRTYEGWDEGHSESLHYEGRAVDITTSDRDRSKYGMARLAVAGFDWYYESKA 180
 DB 121 VRLRTYEGWDEGHSESLHYEGRAVDITTSDRDRSKYGMARLAVAGFDWYYESKA 180
 QY 181 HHCYKAKENSVAASKSGCGFPGSATVHLEOGCTKLVKDLRPGDRVLAADOGRLYSDFL 240
 DB 181 HHCYKAKENSVAASKSGCGFPGSATVHLEOGCTKLVKDLRPGDRVLAADOGRLYSDFL 240
 QY 241 TFLDRDGAKKVFYVETLEPRERLLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 300
 DB 241 TFLDRDGAKKVFYVETLEPRERLLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 300
 QY 301 VYVAERGGDRLLPAAVHSVTLREEAGAYAPLTAHGTLINRVLASCYAVIEESHMAH 360
 DB 301 VYVAERGGDRLLPAAVHSVTLREEAGAYAPLTAHGTLINRVLASCYAVIEESHMAH 360
 QY 361 RAFAFRLHALALAPARTDGGGGSIIPAOSATEARGAEPGTAGIHYSOLLHYHIGTW 420
 DB 361 RAFAFRLHALALAPARTDGGGGSIIPAOSATEARGAEPGTAGIHYSOLLHYHIGTW 420
 QY 421 LLDSETHMPLGMAVKSS 437
 DB 421 LLDSETHMPLGMAVKSS 437

RESULT 2
 US-08-356-060A-11
 ; Sequence 11, Application US/08356060A
 ; Patent No. 5844079
 ; GENERAL INFORMATION:
 ; APPLICANT: Ingham, Phillip W.
 ; APPLICANT: McMahon, Andrew P.
 ; APPLICANT: Tablin, Clifford J.
 ; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
 ; NUMBER OF SEQUENCES: 47
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: ASCII(text)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/356,060A
 ; FILING DATE: 14-DEC-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/176,427
 ; FILING DATE: 30-DEC-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Vincent, Matthew P.
 ; REGISTRATION NUMBER: 36,709
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400
 ; TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 11:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 437 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-356-060A-11

Query Match 100.0%; Score 2289; DB 2; Length 437;
 Best Local Similarity 100.0%; Pred. No. 2, 7e-249;
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLILARCFVLIASSLTCVPGACGRRGRRHKKLPLAYKOFINVAEKTIGAS 60
 DB 1 MLILARCFVLIASSLTCVPGACGRRGRRHKKLPLAYKOFINVAEKTIGAS 60
 QY 61 GYEGKITRNSERFELTPNYNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMWPG 120
 DB 61 GYEGKITRNSERFELTPNYNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMWPG 120
 QY 121 VRLRTYEGWDEGHSESLHYEGRAVDITTSDRDRSKYGMARLAVAGFDWYYESKA 180
 DB 121 VRLRTYEGWDEGHSESLHYEGRAVDITTSDRDRSKYGMARLAVAGFDWYYESKA 180
 QY 181 HHCYKAKENSVAASKSGCGFPGSATVHLEOGCTKLVKDLRPGDRVLAADOGRLYSDFL 240
 DB 181 HHCYKAKENSVAASKSGCGFPGSATVHLEOGCTKLVKDLRPGDRVLAADOGRLYSDFL 240
 QY 241 TFLDRDGAKKVFYVETLEPRERLLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 300
 DB 241 TFLDRDGAKKVFYVETLEPRERLLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 300
 QY 301 VYVAERGGDRLLPAAVHSVTLREEAGAYAPLTAHGTLINRVLASCYAVIEESHMAH 360
 DB 301 VYVAERGGDRLLPAAVHSVTLREEAGAYAPLTAHGTLINRVLASCYAVIEESHMAH 360
 QY 361 RAFAFRLHALALAPARTDGGGGSIIPAOSATEARGAEPGTAGIHYSOLLHYHIGTW 420
 DB 361 RAFAFRLHALALAPARTDGGGGSIIPAOSATEARGAEPGTAGIHYSOLLHYHIGTW 420
 QY 421 LLDSETHMPLGMAVKSS 437
 DB 421 LLDSETHMPLGMAVKSS 437

RESULT 3
 US-08-460-900C-11
 ; Sequence 11, Application US/08460900C
 ; Patent No. 6165747
 ; GENERAL INFORMATION:
 ; APPLICANT: Ingham, Phillip W.
 ; APPLICANT: McMahon, Andrew P.
 ; APPLICANT: Tablin, Clifford J.
 ; APPLICANT: Bumcrot, David A.
 ; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
 ; NUMBER OF SEQUENCES: 62
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 ; STREET: One Post Office Square
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/460,900C
 ; FILING DATE: 5-JUNE-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/435,093
 ; FILING DATE: 4-MAY-1995
 ; APPLICATION NUMBER: US 08/356,060
 ; FILING DATE: 14-DEC-1994
 ; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMV-006.05
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-900C-11

Query Match 100.0%; Score 2289; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 2.7e-249;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLIARCFVLIASSLVCPGLACPGRGFGKRRHPPKLLPLAYKOFIPNVAEKTIGAS 60
DB 1 MLLIARCFVLIASSLVCPGLACPGRGFGKRRHPPKLLPLAYKOFIPNVAEKTIGAS 60
QY 61 GRYEKGITRNSERFELTPNPNPDIIFFKDEENTGADRLMTORCKKLNALAIYVNMOPG 120
DB 61 GRYEKGITRNSERFELTPNPNPDIIFFKDEENTGADRLMTORCKKLNALAIYVNMOPG 120
QY 121 VRLRTTEGDEGHHSESLHIEGRAVDITTSDRRSKYGMARLAVENGFWMYTESKA 180
DB 121 VRLRTTEGDEGHHSESLHIEGRAVDITTSDRRSKYGMARLAVENGFWMYTESKA 180
QY 181 HIHCSVKAENSVAAKSGCFPGSATVHLEOGSTKLVKDLRPDRVLAADDGRLLYSDFL 240
DB 181 HIHCSVKAENSVAAKSGCFPGSATVHLEOGSTKLVKDLRPDRVLAADDGRLLYSDFL 240
QY 241 TFLDDEGAKKVFYIETLEPRERLLTAHLLEFVAPHNDSGPTPGSALFASRYRPGQR 300
DB 241 TFLDDEGAKKVFYIETLEPRERLLTAHLLEFVAPHNDSGPTPGSALFASRYRPGQR 300
QY 301 VYVVAERGGDRRLPAAVHSTLREEAGAVAPLTAHGTILINRVLASCYAVIEESHMAH 360
DB 301 VYVVAERGGDRRLPAAVHSTLREEAGAVAPLTAHGTILINRVLASCYAVIEESHMAH 360
QY 361 RAFAFRLAHALLALAPARTDGGGGSIPAAQSATEARGAEPAGIMHYSQLLYHIGTW 420
DB 361 RAFAFRLAHALLALAPARTDGGGGSIPAAQSATEARGAEPAGIMHYSQLLYHIGTW 420
QY 421 LIDSETMHPGLMAVKSS 437
DB 421 LIDSETMHPGLMAVKSS 437

RESULT 4
US-08-674-509B-11
Sequence 11, Application US/08674509B
Patent No. 6261786
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
APPLICANT: Matigo, Valeria
TITLE OF INVENTION: SCREENING ASSAYS FOR HEDGROG AGONISTS
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/674,509B
FILING DATE: 02-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/460,900
FILING DATE: 05-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMV-006.06
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-674-509B-11

Query Match 100.0%; Score 2289; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 2.7e-249;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLIARCFVLIASSLVCPGLACPGRGFGKRRHPPKLLPLAYKOFIPNVAEKTIGAS 60
DB 1 MLLIARCFVLIASSLVCPGLACPGRGFGKRRHPPKLLPLAYKOFIPNVAEKTIGAS 60
QY 61 GRYEKGITRNSERFELTPNPNPDIIFFKDEENTGADRLMTORCKKLNALAIYVNMOPG 120
DB 61 GRYEKGITRNSERFELTPNPNPDIIFFKDEENTGADRLMTORCKKLNALAIYVNMOPG 120
QY 121 VRLRTTEGDEGHHSESLHIEGRAVDITTSDRRSKYGMARLAVENGFWMYTESKA 180
DB 121 VRLRTTEGDEGHHSESLHIEGRAVDITTSDRRSKYGMARLAVENGFWMYTESKA 180
QY 181 HIHCSVKAENSVAAKSGCFPGSATVHLEOGSTKLVKDLRPDRVLAADDGRLLYSDFL 240
DB 181 HIHCSVKAENSVAAKSGCFPGSATVHLEOGSTKLVKDLRPDRVLAADDGRLLYSDFL 240
QY 241 TFLDDEGAKKVFYIETLEPRERLLTAHLLEFVAPHNDSGPTPGSALFASRYRPGQR 300
DB 241 TFLDDEGAKKVFYIETLEPRERLLTAHLLEFVAPHNDSGPTPGSALFASRYRPGQR 300
QY 301 VYVVAERGGDRRLPAAVHSTLREEAGAVAPLTAHGTILINRVLASCYAVIEESHMAH 360
DB 301 VYVVAERGGDRRLPAAVHSTLREEAGAVAPLTAHGTILINRVLASCYAVIEESHMAH 360
QY 361 RAFAFRLAHALLALAPARTDGGGGSIPAAQSATEARGAEPAGIMHYSQLLYHIGTW 420
DB 361 RAFAFRLAHALLALAPARTDGGGGSIPAAQSATEARGAEPAGIMHYSQLLYHIGTW 420
QY 421 LIDSETMHPGLMAVKSS 437
DB 421 LIDSETMHPGLMAVKSS 437

RESULT 5
US-08-954-698-11
Sequence 11, Application US/08954698
Patent No. 6271363
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing

```

? TITLE OF INVENTION: Proteins and Uses Related Thereto
? NUMBER OF SEQUENCES: 48
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: FOLEY, HOAG & ELIOT LLP
? STREET: One Post Office Square
? CITY: Boston
? STATE: MA
? COUNTRY: USA
? ZIP: 02109-2170
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/954,698
? FILING DATE: 20-OCT-1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/462,386
? FILING DATE: 05-JUN-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/435,093
? FILING DATE: 04-MAY-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/356,060
? FILING DATE: 14-DEC-1994
? APPLICATION NUMBER: US 08/176,427
? FILING DATE: 30-DEC-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Vincent, Matthew P.
? REGISTRATION NUMBER: 36,709
? REFERENCE/DOCKET NUMBER: HMV-006.10
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-832-1000
? TELEFAX: 617-832-7000
? INFORMATION FOR SEQ ID NO: 11:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 437 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-954-698-11

Query Match
Best Local Similarity 100.0%; Score 2289; DB 4; Length 437;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLARCFVLVIASSLVCPCGACGPGRGFGRRRHPKLTPLAYKQFIIPNVAEKTIGAS 60
DB 1 MLLLARCFVLVIASSLVCPCGACGPGRGFGRRRHPKLTPLAYKQFIIPNVAEKTIGAS 60
QY 61 GRYEKGITRNSRFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMQNG 120
DB 61 GRYEKGITRNSRFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMQNG 120
QY 121 VRLRTGEMDEGHHSESLIHYEGNAVDITTSDDRSKYTGMLARLAVAGGDMWYTESKA 180
DB 121 VRLRTGEMDEGHHSESLIHYEGNAVDITTSDDRSKYTGMLARLAVAGGDMWYTESKA 180
QY 181 HIRCSVAENSVAAKSGCGPFGSATVHLEOGCTLVKDLRPGDVLAAADOGRLLYSDFL 240
DB 181 HIRCSVAENSVAAKSGCGPFGSATVHLEOGCTLVKDLRPGDVLAAADOGRLLYSDFL 240
QY 241 TFLDRDGAARKVYVITTEPRERLLTAAHLFFVAPHNDSGPPGSAFASRVPRGQR 300
DB 241 TFLDRDGAARKVYVITTEPRERLLTAAHLFFVAPHNDSGPPGSAFASRVPRGQR 300
QY 301 VYVVAERGGDRLLPLPAVSHVTREEAGAYADLTAGTLLINRVLASCAVVEESHMAH 360
DB 301 VYVVAERGGDRLLPLPAVSHVTREEAGAYADLTAGTLLINRVLASCAVVEESHMAH 360
QY 361 RAFAFRLAHALLAALAPARTDGGCGGSIIPAQOSATFARGAEPYAGIHWY SOLLYHIGTW 420

```

```

DB 361 RAFAFRLAHALLAALAPARTDGGCGGSIIPAQOSATFARGAEPYAGIHWY SOLLYHIGTW 420
QY 421 LIDSETHPLGMVAKSS 437
DB 421 LIDSETHPLGMVAKSS 437

RESULT 6
US-08-957-874-11
? Sequence 11, Application US/08957874
? Patent No. 6384192
? GENERAL INFORMATION:
? APPLICANT: Ingham, Phillip W.
? APPLICANT: McMahon, Andrew P.
? APPLICANT: Tablin, Clifford J.
? TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
? NUMBER OF SEQUENCES: 47
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: FOLEY, HOAG & ELIOT LLP
? STREET: One Post Office Square
? CITY: Boston
? STATE: MA
? COUNTRY: USA
? ZIP: 02109
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: ASCII(text)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/957,874
? FILING DATE: 20-OCT-1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/462,386
? FILING DATE: 5-JUNE-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/435,093
? FILING DATE: 4-MAY-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/356,060
? FILING DATE: 14-DEC-1994
? APPLICATION NUMBER: US 08/176,427
? FILING DATE: 30-DEC-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Vincent, Matthew P.
? REGISTRATION NUMBER: 36,709
? REFERENCE/DOCKET NUMBER: HMV-006.09
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 832-1000
? TELEFAX: (617) 832-7000
? INFORMATION FOR SEQ ID NO: 11:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 437 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-957-874-11

Query Match
Best Local Similarity 100.0%; Score 2289; DB 4; Length 437;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLARCFVLVIASSLVCPCGACGPGRGFGRRRHPKLTPLAYKQFIIPNVAEKTIGAS 60
DB 1 MLLLARCFVLVIASSLVCPCGACGPGRGFGRRRHPKLTPLAYKQFIIPNVAEKTIGAS 60
QY 61 GRYEKGITRNSRFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMQNG 120
DB 61 GRYEKGITRNSRFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMQNG 120

```


Qy	121	VLRAVTEGMDJEDGHHSEESLHYEBRAVDITTSBDRSKYQMLARLVAEAFDMVYYESKA	180
Db	121	VLRAVTEGMDJEDGHHSEESLHYEBRAVDITTSBDRSKYQMLARLVAEAFDMVYYESKA	180
Qy	181	HICSVAEANSVAAKSGCCPGSACTYHLBEOGCKTKLVKDLPRGCRVLAADDOGKILYSDFL	240
Db	181	HICSVAEANSVAAKSGCCPGSACTYHLBEOGCKTKLVKDLPRGCRVLAADDOGKILYSDFL	240
Qy	241	TELDRODGAKVFEVYIETLEPREKRLTLTAANLTFVAPHNDSGPTPGPSALFASRVPRGQR	300
Db	241	TELDRODGAKVFEVYIETLEPREKRLTLTAANLTFVAPHNDSGPTPGPSALFASRVPRGQR	300
Qy	301	VYVVAERGGDRRLPRAVHSTLTLEEBAAGAYAPLTANGTLLINRYLASCAVYEEHSMAH	360
Db	301	VYVVAERGGDRRLPRAVHSTLTLEEBAAGAYAPLTANGTLLINRYLASCAVYEEHSMAH	360
Qy	361	RAFAPEFLAHALLAALAPARTDGGGGSIPPAOSATFARGAEBPTAGIHWYSQLYHIGTW	420
Db	361	RAFAPEFLAHALLAALAPARTDGGGGSIPPAOSATFARGAEBPTAGIHWYSQLYHIGTW	420
Qy	421	LLDSETHMPICMAVKKSS 437	
Db	421	LLDSETHMPICMAVKKSS 437	

```

: RESULT 7
: US-09-325-256-20
: Sequence 20, Application US/09325256
: Patent No. 644793
:
: GENERAL INFORMATION:
: APPLICANT: PEPINSKY, R. BLAKE
: APPLICANT: BAKER, DARREN P.
: APPLICANT: MEN, DINGYI
: APPLICANT: WILLIAMS, KEVIN P.
: APPLICANT: GARGER, ELLEN A.
: APPLICANT: TAYLOR, FREDERICK R.
: APPLICANT: GALDES, ALPHONSE
: APPLICANT: PORTER, JEFFREY
:
: TITLE OF INVENTION: HYDROPHOBICALLY-MODIFIED PROTEIN COMPOSITIONS AND
: FILE REFERENCE: BIV-067.01
:
: CURRENT APPLICATION NUMBER: US/09/325,256
: PRIOR FILING DATE: 1999-06-03
:
: PRIOR APPLICATION NUMBER: 60/099,800
: PRIOR FILING DATE: 1998-09-10
:
: PRIOR APPLICATION NUMBER: 60/078,935
: PRIOR FILING DATE: 1998-03-20
:
: PRIOR APPLICATION NUMBER: 60/089,685
: PRIOR FILING DATE: 1998-06-17
:
: PRIOR APPLICATION NUMBER: 60/067,423
: PRIOR FILING DATE: 1997-12-03
:
: PRIOR APPLICATION NUMBER: PCT/US98/25676
: PRIOR FILING DATE: 1998-12-03
:
: NUMBER OF SEQ ID NOS: 31
:
: SOFTWARE: PatentIn Ver. 2.1
:
: SEQ ID NO 20
:
: LENGTH: 437
:
: TYPE: prt
:
: ORGANISM: Murline sp.
:
: US-09-325-256-20

```

[illegible]

Qy	121	VRLEVTGCMQEDGHHSESLHYEBERADITTSBDRSKYGMRLARLVAEAGFDWYVESKA	180
Db	121	VRLEVTGCMQEDGHHSESLHYEBERADITTSBDRSKYGMRLARLVAEAGFDWYVESKA	180
Qy	181	HHSVSVAEENSVAAKSGGCGPPGSAATYHLEBGGCTKLYVDLARPGRVYLAADQGRILYSDFL	240
Db	181	HHSVSVAEENSVAAKSGGCGPPGSAATYHLEBGGCTKLYVDLARPGRVYLAADQGRILYSDFL	240
Qy	241	TFELDRDGAQKVFYVETLEBRERLLTAAHLTFVAPHNDSGPTPGPSALFASKVRPGQR	300
Db	241	TFELDRDGAQKVFYVETLEBRERLLTAAHLTFVAPHNDSGPTPGPSALFASKVRPGQR	300
Qy	301	VYVAERGGDRRLLPAAVHSTYLEEBEAGAAPLTLAGTLLINRYLASCAVYIEESHMAH	360
Db	301	VYVAERGGDRRLLPAAVHSTYLEEBEAGAAPLTLAGTLLINRYLASCAVYIEESHMAH	360
Qy	361	RAFAFRLAHLAALAIPARTDGGGGSIPAAOSATEARGAEPYAGIMHYSOLLYHIGTW	420
Db	361	RAFAFRLAHLAALAIPARTDGGGGSIPAAOSATEARGAEPYAGIMHYSOLLYHIGTW	420
Qy	421	LLDSETHMPICMAVKSS 437	
Db	421	LLDSETHMPICMAVKSS 437	

```

US-08-946-329A-20
SEQUENCE 20, APPLICATION US/08946329A
PATENT NO. 6057091
GENERAL INFORMATION:
APPLICANT: Beachy, Philip A.
APPLICANT: Porter, Jeffrey A.
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED P
NUMBER OF SEQUENCES: 109
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0.bb
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/946,329A
FILING DATE: 07-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/061,323
FILING DATE: 07-OCT-1996
APPLICATION NUMBER: 08/729,743
FILING DATE: 10-JUL-1996
APPLICATION NUMBER: 08/567,357
FILING DATE: 04-DEC-1995
APPLICATION NUMBER: 08/349,498
FILING DATE: 02-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/140001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-946-329A-20

```

Query Match 99.9%; Score 2286; DB 3; Length 437;
 Best Local Similarity 99.8%; Pred. No. 5,9e-249;
 Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MLLLLARCFVLIIASSLLVCPGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIIASSLLVCPGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
OY 61 GRYEGKTRNSERKELTPNPNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQMPG 120
DB 61 GRYEGKTRNSERKELTPNPNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQMPG 120
OY 121 VRLRVTEGMDEDEGHSEESLHYEGRAVDITTSDRDSKYGMARLAVEAGFDMVYTESKA 180
DB 121 VRLRVTEGMDEDEGHSEESLHYEGRAVDITTSDRDSKYGMARLAVEAGFDMVYTESKA 180
OY 181 HHCISVKAENSVAASKSGCFPGSATVHLEOGGKTVKDLRPDRLVLAADOGRLIYSDFL 240
DB 181 HHCISVKAENSVAASKSGCFPGSATVHLEOGGKTVKDLRPDRLVLAADOGRLIYSDFL 240
OY 241 TFLDRREGAKKVFYIETLEPRERLLTAAHLFVAPHNDSGPTPGSALFASRVPRGQR 300
DB 241 TFLDRREGAKKVFYIETLEPRERLLTAAHLFVAPHNDSGPTPGSALFASRVPRGQR 300
OY 301 VYVVAERGGDRLLPAAVHSTVTLREEGAGAYAPLTAGTILINRVLASCYAVIEESHMAH 360
DB 301 VYVVAERGGDRLLPAAVHSTVTLREEGAGAYAPLTAGTILINRVLASCYAVIEESHMAH 360
OY 361 RAFAFRILAHALLAALAPARTDGGGGSIPAOSATFARGAEPPTAGIHWYSOLLHYHIGTW 420
DB 361 RAFAFRILAHALLAALAPARTDGGGGSIPAOSATFARGAEPPTAGIHWYSOLLHYHIGTW 420
OY 421 LLDSETHMPLGMAVKSS 437
DB 421 LLDSETHMPLGMAVKSS 437

```

RESULT 9

US-08-567-357A-20
 ; Sequence 20, Application US/08567357A
 ; Patent No. 6132728
 ; GENERAL INFORMATION:
 ; APPLICANT: Beachy, Phillip A.
 ; APPLICANT: Moon, Randall T.
 ; APPLICANT: Porter, Jeffrey A.
 ; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 37
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 4225 Executive Square, Suite 1400
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows 95
 ; SOFTWARE: FASTSEQ for Windows Version 2.0b
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/567,357A
 ; FILING DATE: 04-DEC-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/349,498
 ; FILING DATE: 02-DEC-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Halle, Lisa A.
 ; REGISTRATION NUMBER: 38,347
 ; REFERENCE/DOCKET NUMBER: 07265/080001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619/678-5070
 ; TELEFAX: 619/678-5099
 ; INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:
 ; LENGTH: 437 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-567-357A-20

Query Match 99.9%; Score 2286; DB 4; Length 437;
 Best Local Similarity 99.8%; Pred. No. 5,9e-249;
 Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MLLLLARCFVLIIASSLLVCPGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIIASSLLVCPGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
OY 61 GRYEGKTRNSERKELTPNPNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQMPG 120
DB 61 GRYEGKTRNSERKELTPNPNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQMPG 120
OY 121 VRLRVTEGMDEDEGHSEESLHYEGRAVDITTSDRDSKYGMARLAVEAGFDMVYTESKA 180
DB 121 VRLRVTEGMDEDEGHSEESLHYEGRAVDITTSDRDSKYGMARLAVEAGFDMVYTESKA 180
OY 181 HHCISVKAENSVAASKSGCFPGSATVHLEOGGKTVKDLRPDRLVLAADOGRLIYSDFL 240
DB 181 HHCISVKAENSVAASKSGCFPGSATVHLEOGGKTVKDLRPDRLVLAADOGRLIYSDFL 240
OY 241 TFLDRREGAKKVFYIETLEPRERLLTAAHLFVAPHNDSGPTPGSALFASRVPRGQR 300
DB 241 TFLDRREGAKKVFYIETLEPRERLLTAAHLFVAPHNDSGPTPGSALFASRVPRGQR 300
OY 301 VYVVAERGGDRLLPAAVHSTVTLREEGAGAYAPLTAGTILINRVLASCYAVIEESHMAH 360
DB 301 VYVVAERGGDRLLPAAVHSTVTLREEGAGAYAPLTAGTILINRVLASCYAVIEESHMAH 360
OY 361 RAFAFRILAHALLAALAPARTDGGGGSIPAOSATFARGAEPPTAGIHWYSOLLHYHIGTW 420
DB 361 RAFAFRILAHALLAALAPARTDGGGGSIPAOSATFARGAEPPTAGIHWYSOLLHYHIGTW 420
OY 421 LLDSETHMPLGMAVKSS 437
DB 421 LLDSETHMPLGMAVKSS 437

```

RESULT 10

US-08-729-743A-20
 ; Sequence 20, Application US/08729743A
 ; Patent No. 6214794
 ; GENERAL INFORMATION:
 ; APPLICANT: Beachy, Phillip A.
 ; APPLICANT: Moon, Randall T.
 ; APPLICANT: Porter, Jeffrey A.
 ; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 37
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 4225 Executive Square, Suite 1400
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows 95
 ; SOFTWARE: FASTSEQ for Windows Version 2.0b
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/729,743A
 ; FILING DATE: 07-OCT-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/567,357
 ; FILING DATE: 04-DEC-1995
 ; APPLICATION NUMBER: 08/349,498

FILING DATE: 02-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/099001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-729-743A-20

Query Match 99.9%; Score 2286; DB 4; Length 437;
Best Local Similarity 99.8%; Pred. No. 5.9e-249;
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLARCFVLVIASSLVCPGLACPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLARCFVLVIASSLVCPGLACPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
QY 61 GRYEKGITRNSERFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAIISVNQWPG 120
DB 61 GRYEKGITRNSERFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAIISVNQWPG 120
QY 121 VRLRTGEGDEGHSESLHSEGRAVDITTSDRDSKYGMLARLAVEGFPMVYTESKA 180
DB 121 VRLRTGEGDEGHSESLHSEGRAVDITTSDRDSKYGMLARLAVEGFPMVYTESKA 180
QY 181 HIHCVKAENSVAAKSGCGFPGSATVHLEQGTGKLVKDLRPGDRVLAADDOGRLLYSDFL 240
DB 181 HIHCVKAENSVAAKSGCGFPGSATVHLEQGTGKLVKDLRPGDRVLAADDOGRLLYSDFL 240
QY 241 TELDRDEGAKKVFYIETLEPERELLTAHLLFVAPHNDSGPTPGPSALFASRVPPGQR 300
DB 241 TELDRDEGAKKVFYIETLEPERELLTAHLLFVAPHNDSGPTPGPSALFASRVPPGQR 300
QY 301 VYVVAERGGDRLLPAVAHVSITLREEAGAYAPLTAHGTLINRVLASCYAVIEEHSMAH 360
DB 301 VYVVAERGGDRLLPAVAHVSITLREEAGAYAPLTAHGTLINRVLASCYAVIEEHSMAH 360
QY 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWSQLVHIGTW 420
DB 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWSQLVHIGTW 420
QY 421 LIDSETMHPILGMAVKSS 437
DB 421 LIDSETMHPILGMAVKSS 437

RESULT 11
US-09-057-860A-6
Sequence 6, Application US/09057860A
Patent No. 6277820
GENERAL INFORMATION:
APPLICANT: Arnon Rosenthal
APPLICANT: Mary Hynes
APPLICANT: Weilan Ye
TITLE OF INVENTION: Method Of Dopaminergic And Seratonegic
TITLE OF INVENTION: Neuron Formation From Neuroprogenitor Cells
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057, 860A
FILING DATE: 09-Apr-1998
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Svoboda, Craig G.
REGISTRATION NUMBER: 39,044
REFERENCE/DOCKET NUMBER: P1364
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1489
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: Amino Acid
TOPOLOGY: linear
US-09-057-860A-6

Query Match 99.9%; Score 2286; DB 4; Length 437;
Best Local Similarity 99.8%; Pred. No. 5.9e-249;
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLARCFVLVIASSLVCPGLACPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLARCFVLVIASSLVCPGLACPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
QY 61 GRYEKGITRNSERFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAIISVNQWPG 120
DB 61 GRYEKGITRNSERFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAIISVNQWPG 120
QY 121 VRLRTGEGDEGHSESLHSEGRAVDITTSDRDSKYGMLARLAVEGFPMVYTESKA 180
DB 121 VRLRTGEGDEGHSESLHSEGRAVDITTSDRDSKYGMLARLAVEGFPMVYTESKA 180
QY 181 HIHCVKAENSVAAKSGCGFPGSATVHLEQGTGKLVKDLRPGDRVLAADDOGRLLYSDFL 240
DB 181 HIHCVKAENSVAAKSGCGFPGSATVHLEQGTGKLVKDLRPGDRVLAADDOGRLLYSDFL 240
QY 241 TELDRDEGAKKVFYIETLEPERELLTAHLLFVAPHNDSGPTPGPSALFASRVPPGQR 300
DB 241 TELDRDEGAKKVFYIETLEPERELLTAHLLFVAPHNDSGPTPGPSALFASRVPPGQR 300
QY 301 VYVVAERGGDRLLPAVAHVSITLREEAGAYAPLTAHGTLINRVLASCYAVIEEHSMAH 360
DB 301 VYVVAERGGDRLLPAVAHVSITLREEAGAYAPLTAHGTLINRVLASCYAVIEEHSMAH 360
QY 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWSQLVHIGTW 420
DB 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWSQLVHIGTW 420
QY 421 LIDSETMHPILGMAVKSS 437
DB 421 LIDSETMHPILGMAVKSS 437

RESULT 12
US-08-349-498-20
Sequence 20, Application US/08349498
Patent No. 6281332
GENERAL INFORMATION:
APPLICANT: Beachy, Phillip A.
APPLICANT: Moon, Randall T.
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: LA Jolla
STATE: CA
COUNTRY: U.S.A.

```

; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/349,498
; FILING DATE: 02-DEC-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/043001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-349-498-20

```

```

Query Match          99.9%; Score 2286; DB 4; Length 437;
Best Local Similarity 99.8%; Pred. No. 5.9e-249;
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLIARCFVLIASSLVCPGLACGPGRGFRGRRPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLIARCFVLIASSLVCPGLACGPGRGFRGRRPKLTPLAYKQFIPNVAEKTIGAS 60
QY 61 GREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALATSYMNQWPG 120
DB 61 GREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALATSYMNQWPG 120
QY 121 VRLRVTEGWDEDEGHSESLHTEGRAVDITTSRDRSKYGMRLARLAVEGFDMVYESKA 180
DB 121 VRLRVTEGWDEDEGHSESLHTEGRAVDITTSRDRSKYGMRLARLAVEGFDMVYESKA 180
QY 181 HHCYSKAKNSVAAKSGGCPGSAIVHLEGGTKLVKDLRGDRVLAADDGGRLLYSDFL 240
DB 181 HHCYSKAKNSVAAKSGGCPGSAIVHLEGGTKLVKDLRGDRVLAADDGGRLLYSDFL 240
QY 241 TFLDDEGAKKVFYIETLEPERILLTAAHLFVAAPHNDGPTPGSALFASVRPGROR 300
DB 241 TFLDDEGAKKVFYIETLEPERILLTAAHLFVAAPHNDGPTPGSALFASVRPGROR 300
QY 301 VYVAERGGDRRLPAVAHSTVTLREEEAGAYAPLTAHGTILLINRVLASCYAVIEESHMAH 360
DB 301 VYVAERGGDRRLPAVAHSTVTLREEEAGAYAPLTAHGTILLINRVLASCYAVIEESHMAH 360
QY 361 RAFAFRLAHLAALAPARTDGGGSIIPAASATBARGAETAGIHWYSOLLHYHIGTW 420
DB 361 RAFAFRLAHLAALAPARTDGGGSIIPAASATBARGAETAGIHWYSOLLHYHIGTW 420
QY 421 LLDSETHMPLGMAVKSS 437
DB 421 LLDSETHMPLGMAVKSS 437

```

```

RESULT 13
US-09-293-505-14
; Sequence 14, Application US/09293505
; Patent No. 6348575
; GENERAL INFORMATION:
; APPLICANT: de Sauvage, Frederic
; TITLE OF INVENTION: Patched-2
; FILE REFERENCE: P1405R1
; CURRENT APPLICATION NUMBER: US/09/293,505
; CURRENT FILING DATE: 1999-04-15

```

```

; EARLIER APPLICATION NUMBER: US 60/081,884
; EARLIER FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 14
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-293-505-14

```

```

Query Match          99.9%; Score 2286; DB 4; Length 437;
Best Local Similarity 99.8%; Pred. No. 5.9e-249;
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLIARCFVLIASSLVCPGLACGPGRGFRGRRPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLIARCFVLIASSLVCPGLACGPGRGFRGRRPKLTPLAYKQFIPNVAEKTIGAS 60
QY 61 GREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALATSYMNQWPG 120
DB 61 GREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALATSYMNQWPG 120
QY 121 VRLRVTEGWDEDEGHSESLHTEGRAVDITTSRDRSKYGMRLARLAVEGFDMVYESKA 180
DB 121 VRLRVTEGWDEDEGHSESLHTEGRAVDITTSRDRSKYGMRLARLAVEGFDMVYESKA 180
QY 181 HHCYSKAKNSVAAKSGGCPGSAIVHLEGGTKLVKDLRGDRVLAADDGGRLLYSDFL 240
DB 181 HHCYSKAKNSVAAKSGGCPGSAIVHLEGGTKLVKDLRGDRVLAADDGGRLLYSDFL 240
QY 241 TFLDDEGAKKVFYIETLEPERILLTAAHLFVAAPHNDGPTPGSALFASVRPGROR 300
DB 241 TFLDDEGAKKVFYIETLEPERILLTAAHLFVAAPHNDGPTPGSALFASVRPGROR 300
QY 301 VYVAERGGDRRLPAVAHSTVTLREEEAGAYAPLTAHGTILLINRVLASCYAVIEESHMAH 360
DB 301 VYVAERGGDRRLPAVAHSTVTLREEEAGAYAPLTAHGTILLINRVLASCYAVIEESHMAH 360
QY 361 RAFAFRLAHLAALAPARTDGGGSIIPAASATBARGAETAGIHWYSOLLHYHIGTW 420
DB 361 RAFAFRLAHLAALAPARTDGGGSIIPAASATBARGAETAGIHWYSOLLHYHIGTW 420
QY 421 LLDSETHMPLGMAVKSS 437
DB 421 LLDSETHMPLGMAVKSS 437

```

```

RESULT 14
PCT-US95-15463-20
; Sequence 20, Application PC/TUS9515463
; GENERAL INFORMATION:
; APPLICANT: The Johns Hopkins University School of Medicine
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15463
; FILING DATE: 01-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/080W01

```

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619/678-5070
 TELEFAX: 619/678-5099
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 437 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US95-15463-20

Query Match 99.9%; Score 2286; DB 5; Length 437;
 Best Local Similarity 99.8%; Pred. No. 5.9e-249;
 Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLLIARCFVLIYLLSSLYCPGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
 DB 1 MLLIARCFVLIYLLSSLYCPGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
 OY 61 GRYEGKITRNSERFKELTPNPNPDIIFFKDEENTGADRLMTQRCCKLNALAISVNMWPG 120
 DB 61 GRYEGKITRNSERFKELTPNPNPDIIFFKDEENTGADRLMTQRCCKLNALAISVNMWPG 120
 OY 121 VRLVTEGMDGDGHHSESLHYEGRAVDITTSDRRSKYGMRLARLAVAGFDWYTESKA 180
 DB 121 VRLVTEGMDGDGHHSESLHYEGRAVDITTSDRRSKYGMRLARLAVAGFDWYTESKA 180
 OY 181 HHCSSVAENSVAAKSGCGFPGSATVHLEGGTKLVKDLRPGDRVLAADDGRLLYSDFL 240
 DB 181 HHCSSVAENSVAAKSGCGFPGSATVHLEGGTKLVKDLRPGDRVLAADDGRLLYSDFL 240
 OY 241 TFLDDEGAKKVFYIETLEPERELLTPAHLLEFVAPHNDSGPTPGSPALFASRYRPGQR 300
 DB 241 TFLDDEGAKKVFYIETLEPERELLTPAHLLEFVAPHNDSGPTPGSPALFASRYRPGQR 300
 OY 301 VYVVAERGGDRRLPAAVHSTLREEAGAYAPLTAHGTIILINRYLASCYAVIEESHNAH 360
 DB 301 VYVVAERGGDRRLPAAVHSTLREEAGAYAPLTAHGTIILINRYLASCYAVIEESHNAH 360
 OY 361 RAFAFRLAHLALLAALAPARTDGGGGGSIIPAQSATEARGAEPAGIHWSOLLYHIGTW 420
 DB 361 RAFAFRLAHLALLAALAPARTDGGGGGSIIPAQSATEARGAEPAGIHWSOLLYHIGTW 420
 OY 421 LLDSETHPLGMAVKSS 437
 DB 421 LLDSETHPLGMAVKSS 437

RESULT 15
 PCT-US95-15923-20
 Sequence 20, Application PC/TUS9515923
 GENERAL INFORMATION:
 APPLICANT: The Johns Hopkins University School of Medicine, et al.
 TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 4225 Executive Square, Suite 1400
 CITY: La Jolla
 STATE: CA
 COUNTRY: U.S.A.
 ZIP: 92037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentln Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/15923
 FILING DATE: 04-DEC-1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Halle, Lisa A.

REGISTRATION NUMBER: 38,347
 REFERENCE/DOCKET NUMBER: 07265/043W01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619/678-5070
 TELEFAX: 619/678-5099
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 437 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US95-15923-20

Query Match 99.9%; Score 2286; DB 5; Length 437;
 Best Local Similarity 99.8%; Pred. No. 5.9e-249;
 Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLLIARCFVLIYLLSSLYCPGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
 DB 1 MLLIARCFVLIYLLSSLYCPGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
 OY 61 GRYEGKITRNSERFKELTPNPNPDIIFFKDEENTGADRLMTQRCCKLNALAISVNMWPG 120
 DB 61 GRYEGKITRNSERFKELTPNPNPDIIFFKDEENTGADRLMTQRCCKLNALAISVNMWPG 120
 OY 121 VRLVTEGMDGDGHHSESLHYEGRAVDITTSDRRSKYGMRLARLAVAGFDWYTESKA 180
 DB 121 VRLVTEGMDGDGHHSESLHYEGRAVDITTSDRRSKYGMRLARLAVAGFDWYTESKA 180
 OY 181 HHCSSVAENSVAAKSGCGFPGSATVHLEGGTKLVKDLRPGDRVLAADDGRLLYSDFL 240
 DB 181 HHCSSVAENSVAAKSGCGFPGSATVHLEGGTKLVKDLRPGDRVLAADDGRLLYSDFL 240
 OY 241 TFLDDEGAKKVFYIETLEPERELLTPAHLLEFVAPHNDSGPTPGSPALFASRYRPGQR 300
 DB 241 TFLDDEGAKKVFYIETLEPERELLTPAHLLEFVAPHNDSGPTPGSPALFASRYRPGQR 300
 OY 301 VYVVAERGGDRRLPAAVHSTLREEAGAYAPLTAHGTIILINRYLASCYAVIEESHNAH 360
 DB 301 VYVVAERGGDRRLPAAVHSTLREEAGAYAPLTAHGTIILINRYLASCYAVIEESHNAH 360
 OY 361 RAFAFRLAHLALLAALAPARTDGGGGGSIIPAQSATEARGAEPAGIHWSOLLYHIGTW 420
 DB 361 RAFAFRLAHLALLAALAPARTDGGGGGSIIPAQSATEARGAEPAGIHWSOLLYHIGTW 420
 OY 421 LLDSETHPLGMAVKSS 437
 DB 421 LLDSETHPLGMAVKSS 437

Search completed: February 20, 2003, 10:14:11
 Job time : 12.4398 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:08:18 ; Search time 14.3815 seconds
(without alignments)
2921.176 Million cell updates/sec

Title: US-09-827-110a-13

Perfect score: 2289

Sequence: 1 MLLILARCFVLIASSLLVC.....GTWLDSETHMPLGNVAKSS 437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2286	99.9	437	2 A49425	Sonic hedgehog pro
2	2200	96.1	437	2 B53193	hedgehog homolog v
3	1788	78.1	425	2 A49424	patterning protein
4	1525.5	66.6	444	2 S56765	morphogen Xhh prec
5	1517	66.3	418	2 A53193	hedgehog homolog v
6	1458	63.7	415	2 A49426	sonic hedgehog gen
7	1219	53.3	396	2 B49425	Desert hedgehog pr
8	1043	45.6	336	2 C49425	Indian hedgehog pr
9	894	39.1	471	2 A46400	segment polarity p
10	452	19.7	94	2 G02735	desert hedgehog -
11	170	7.4	615	2 T29550	hypothetical prote
12	153.5	6.7	1207	2 T23754	hypothetical prote
13	148	6.5	1226	2 T24045	hypothetical prote
14	146	6.4	481	2 T27655	hypothetical prote
15	140.5	6.1	629	2 T19563	hypothetical prote
16	130.5	5.7	868	2 T22281	hypothetical prote
17	128	5.6	1021	2 T23252	hypothetical prote
18	127	5.5	484	2 T34504	hypothetical prote
19	118	5.2	598	2 T42070	protein serine/thr
20	117	5.1	1407	2 T28702	probable polyketid
21	108	4.7	481	2 T27975	hypothetical prote
22	107	4.7	672	2 T36083	hypothetical prote
23	105	4.6	1324	2 T14070	peptide synthetase
24	102.5	4.5	384	2 H87019	probable glycosyl
25	101	4.4	482	2 E70554	hypothetical prote
26	100.5	4.4	863	2 H84177	mismatch repair pr
27	100.5	4.4	930	2 T35180	hypothetical prote
28	100	4.4	831	2 H84368	MCM / cell divisio
29	100	4.4	1147	2 D87295	smc protein [impor

30	99	4.3	869	2 S76720	hypothetical prote
31	98.5	4.3	532	2 E87343	ABC transporter, A
32	98.5	4.3	1464	2 T13716	bazooka gene prote
33	98	4.3	940	2 E87250	[protein-pili] urid
34	97.5	4.3	313	2 T15855	hypothetical prote
35	96	4.2	205	2 T26220	hypothetical prote
36	96	4.2	373	2 D36067	thyroid hormone re
37	96	4.2	419	2 AF1274	retrovirus-related
38	96	4.2	572	1 FOHY1H	copper resistance
39	96	4.2	611	2 C82845	probable exoribonu
40	96	4.2	759	2 G71279	hypothetical prote
41	96	4.2	1958	2 B40505	probable outer mem
42	95.5	4.2	1531	2 H71468	gene c-erbA-beta p
43	95	4.2	373	2 T51165	probable anthranil
44	94.5	4.1	335	2 T36304	propiomyl-CoA card
45	94.5	4.1	669	2 B87519	

ALIGNMENTS

RESULT 1

A49425 Sonic hedgehog protein precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999

C:Accession: A49425

R:Reheland, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; M

Cell 75, 1417-1430, 1993

A:Title: Sonic hedgehog, a member of a family of putative signalling molecules, is imp

A:Reference number: A49425; MUID:94094334; PMID:7916661

A:Accession: A49425

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1437 <ECH>

A:Cross-references: GB:X76290

A:Note: authors translation is shown for the codon TCC at position 436

C:Genetics:

A:Gene: Shh

C:Superfamily: sonic hedgehog protein

Query Match		99.9%	Score 2286	DB 2	Length 437
Best Local Similarity		99.8%	Pred. No. 7.2e-177		
Matches 436		Conservative 1	Mismatches 0	Indels 0	Gaps 0
QY	1	MLLILARCFVLIASSLLVC	PGCGRGRCGRKRRPKKLTPLAYKQFI	PVNAEKTIGAS	60
Db	1	MLLILARCFVLIASSLLVC	PGCGRGRCGRKRRPKKLTPLAYKQFI	PVNAEKTIGAS	60
QY	61	GRYGGKTRNSRREKELTPNYNDIIFKDEENTGADRL	TORCKDKLNALAI	SVNOMP	120
Db	61	GRYGGKTRNSRREKELTPNYNDIIFKDEENTGADRL	TORCKDKLNALAI	SVNOMP	120
QY	121	VRLEVTGEWDEGHSESLHYEGRAVDITTSDBRSK	YGMRLAVEAGFDV	YESKA	180
Db	121	VRLEVTGEWDEGHSESLHYEGRAVDITTSDBRSK	YGMRLAVEAGFDV	YESKA	180
QY	181	HIRHSVAENSVAAKSGGCPGSGATVHLEBOGCT	LVNDLRPGDVL	LAADOGRL	YSDFL 240
Db	181	HIRHSVAENSVAAKSGGCPGSGATVHLEBOGCT	LVNDLRPGDVL	LAADOGRL	YSDFL 240
QY	241	TFILDRDGAKKVFYIETLEPRERLLTLAAHLFVA	RHNSGPTPGS	ALFASVRPGR	300
Db	241	TFILDRDGAKKVFYIETLEPRERLLTLAAHLFVA	RHNSGPTPGS	ALFASVRPGR	300
QY	301	VYVVAERGGDRLLPAVHSTVTLREEDAGAVAPLTAGTIL	INNVLASCYA	VEEHSMAH	360
Db	301	VYVVAERGGDRLLPAVHSTVTLREEDAGAVAPLTAGTIL	INNVLASCYA	VEEHSMAH	360
QY	361	RAFAFPLAHALLAALAPARTDGGGGSIPAAGSATRA	GAEPAGIHW	SOLLVYHGTW	420
Db	361	RAFAFPLAHALLAALAPARTDGGGGSIPAAGSATRA	GAEPAGIHW	SOLLVYHGTW	420

QY 421 LLDSETHPLGMAVKSS 437
 |||||
 Db 421 LLDSETHPLGMAVKSS 437

RESULT 2

Best Local Similarity 96.1%; Score 2200; DB 2; Length 437;
 Matches 421; Conservative 3; Mismatches 13; Indels 0; Gaps 0;
 hedgehog homolog vhh-1 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999
 C:Accession: B53193
 R:Roelink, H.; Augsburg, A.; Heemskerk, J.; Korzh, V.; Norlin, S.; Ruiz i Altaba, A.;
 Cell 76, 761-775, 1994
 A:Title: Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedgeh
 A:Reference number: A53193; MUID:94170375; PMID:8124714
 A:Accession: B53193
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-437 <ROE>
 A:Cross-references: GB:L27340; NID:9452122; PIDN:AAA20999.1; PID:9452123
 C:Superfamily: sonic hedgehog protein

Query Match 96.1%; Score 2200; DB 2; Length 437;
 Best Local Similarity 96.3%; Pred. No. 6.5e-170;
 Matches 421; Conservative 3; Mismatches 13; Indels 0; Gaps 0;
 QY 1 MLLLAFCFLVLIASSLVCPGLACGPGRGFRHHPKRLTPLAYKQFIPNVAEKTIGAS 60
 |||||
 Db 1 MLLLAFCFLVLIASSLVCPGLACGPGRGFRHHPKRLTPLAYKQFIPNVAEKTIGAS 60
 QY 61 GYEGKITNSERFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAIISVNQWPG 120
 |||||
 Db 61 GYEGKITNSERFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAIISVNQWPG 120
 QY 121 VRLRTGDEDEGHSEESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYVESKA 180
 |||||
 Db 121 VRLRTGDEDEGHSEESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYVESKA 180
 QY 121 VRLRTGDEDEGHSEESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYVESKA 180
 |||||
 Db 121 VRLRTGDEDEGHSEESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYVESKA 180
 QY 181 HHCYKVAENSVAAKSGCGPSATVHLEOGTKLVKDLRPDRLAADDOGRLLYSDFL 240
 |||||
 Db 181 HHCYKVAENSVAAKSGCGPSATVHLEOGTKLVKDLRPDRLAADDOGRLLYSDFL 240
 QY 241 TFLDRREGAKKVVYETLEPRERLLITAAHLFVAPRHNDGTPPSPALFASRVPRGQR 300
 |||||
 Db 241 TFLDRREGAKKVVYETLEPRERLLITAAHLFVAPRHNDGTPPSPALFASRVPRGQR 300
 QY 301 VVVAERGGDRRLPAVAHSVTLREBEAGAVAPLTAHGTILINRYLASCYAYIEHSMH 360
 |||||
 Db 301 VVVAERGGDRRLPAVAHSVTLREBEAGAVAPLTAHGTILINRYLASCYAYIEHSMH 360
 QY 361 RAFAFRLAHLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWYSOLLVHIGTW 420
 |||||
 Db 361 RAFAFRLAHLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWYSOLLVHIGTW 420
 QY 421 LLDSETHPLGMAVKSS 437
 |||||
 Db 421 LLDSETHPLGMAVKSS 437

RESULT 3

Best Local Similarity 67.4%; Pred. No. 2.1e-115;
 Matches 304; Conservative 38; Mismatches 88; Indels 21; Gaps 6;
 patterning protein Sonic hedgehog precursor - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
 C:Accession: A49424
 R:Riddle, R.D.; Johnson, R.L.; Lauffer, E.; Tabin, C.
 Cell 75, 1401-1416, 1993
 A:Title: Sonic hedgehog mediates the polarizing activity of the ZPA.
 A:Reference number: A49424; MUID:94094333; PMID:8269518
 A:Accession: A49424
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-425 <RID>

A:Cross-references: GB:L28099; NID:9453526; PIDN:AAA72428.1; PID:9453527
 C:Superfamily: sonic hedgehog protein
 F:1-26/Domain: signal sequence #status predicted <SIG>

Query Match 78.1%; Score 1788; DB 2; Length 425;
 Best Local Similarity 80.2%; Pred. No. 1.2e-136;
 Matches 353; Conservative 22; Mismatches 43; Indels 22; Gaps 6;

QY 2 MLLLAFCFLVLIASSLVCPGLACGPGRGFRHHPKRLTPLAYKQFIPNVAEKTIGAS 61
 |||||
 Db 4 MLLLAFCFLVLIASSLVCPGLACGPGRGFRHHPKRLTPLAYKQFIPNVAEKTIGAS 63
 QY 62 RYEGKITNSERFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAIISVNQWPG 121
 |||||
 Db 62 RYEGKITNSERFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAIISVNQWPG 123
 QY 122 RLRTGDEDEGHSEESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYVESKA 181
 |||||
 Db 122 RLRTGDEDEGHSEESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYVESKA 183
 QY 182 HHCYKVAENSVAAKSGCGPSATVHLEOGTKLVKDLRPDRLAADDOGRLLYSDFL 241
 |||||
 Db 182 HHCYKVAENSVAAKSGCGPSATVHLEOGTKLVKDLRPDRLAADDOGRLLYSDFL 243
 QY 242 TFLDRREGAKKVVYETLEPRERLLITAAHLFVAPRHNDGTPPSPALFASRVPR 297
 |||||
 Db 242 TFLDRREGAKKVVYETLEPRERLLITAAHLFVAPRHNDGTPPSPALFASRVPR 302
 QY 298 GGRVYVAERGGDRRLPAVAHSVTLREBEAGAVAPLTAHGTILINRYLASCYAYIEH 357
 |||||
 Db 303 GGRVYVAERGGDRRLPAVAHSVTLREBEAGAVAPLTAHGTILINRYLASCYAYIEH 360
 QY 358 WAHRAFAFRLAHLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWYSOLLVH 417
 |||||
 Db 361 WAHRAFAFRLAHLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWYSOLLVH 405
 QY 418 GTWLDDSETHPLGMAVKSS 437
 |||||
 Db 406 GSWVLDGDLHPLGMAVPAS 425

RESULT 4

Best Local Similarity 66.6%; Score 1525.5; DB 2; Length 444;
 Matches 304; Conservative 38; Mismatches 88; Indels 21; Gaps 6;
 morphogen Xhh precursor - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000
 C:Accession: S56765
 R:Stolow, M.A.; Shi, Y.B.
 Nucleic Acids Res. 23, 2555-2562, 1995
 A:Title: Xenopus sonic hedgehog as a potential morphogen during embryogenesis and thy
 A:Reference number: S56765; MUID:95357169; PMID:7630736
 A:Accession: S56765
 A:Status: preliminary
 A:Molecule type: nucleic acid sequence not shown
 A:Residues: 1-444 <STD>
 A:Cross-references: EMBL:L39213; NID:9790937; PIDN:AA42227.1; PID:9790938
 C:Superfamily: sonic hedgehog protein

Query Match 66.6%; Score 1525.5; DB 2; Length 444;
 Best Local Similarity 67.4%; Pred. No. 2.1e-115;
 Matches 304; Conservative 38; Mismatches 88; Indels 21; Gaps 6;
 QY 1 MLLLAFCFLVLIASSLVCPGLACGPGRGFRHHPKRLTPLAYKQFIPNVAEKTIGAS 60
 |||||
 Db 1 MLLLAFCFLVLIASSLVCPGLACGPGRGFRHHPKRLTPLAYKQFIPNVAEKTIGAS 60
 QY 61 GYEGKITNSERFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAIISVNQWPG 120
 |||||
 Db 61 GYEGKITNSERFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAIISVNQWPG 120
 QY 121 VRLRTGDEDEGHSEESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYVESKA 180
 |||||
 Db 121 VRLRTGDEDEGHSEESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYVESKA 180


```

OY 181 HHCSSVKAENSVAAKSGCGFPGSATVHLEBOGGTKLVKDLRPGDRVLAADOGRLTSDFL 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 HHCSSVKAENSVAAKSGCGFPGACARVMEFGGKAVKDLRPGDRVLSDDPGMLTSDFL 240
OY 241 TFLDRDEGAKKVFYIETLEPERELLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 241 MEIDDERVKKTLFYYIET--SQRKRLTAHLLFVAQTQVNG-TRSFKSVFASNIQPGDL 297
OY 301 VVVAERGGDRRLPAAVHSVTLREBEAGAVAPLTAGHTILINRVLASCYAVIEEHSMAH 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 298 ITTADPK--TWTLKAVKVRKVDL-EDDTAAYAPLTAGHTIVVDQVLASCYAVIEEHTMAH 354
OY 361 RAFAPERLAHLAALAPARTDGGGGGSIPTAAO-----SATEARGAEPTAG 406
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 355 LAFAPLRFGMSLSSYIYR-RDSSPRSGLOPHNOVDQSHNOVDLSHQVLDQSHNOQLGS 413
OY 407 IHMYSQLYHIGTWLSDSETMHPGLMAVSS 437
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 414 IHMYSQLYQIGTWLSDSNSLHPGLMAVTKSS 444

```

RESULT 5

```

A:3193
hedghehog homolog vhh-1 - zebra fish
C:Species: Brachydanio rerio (zebra fish)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999
C:Accession: A53193
R:Roelink, H.; Augsburg, A.; Heemskerk, J.; Korzh, V.; Norlin, S.; Ruiz i Altaba, A.;
Cell 76, 761-775, 1994
A:Title: Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedgehog
A:Reference number: A53193; MUID:94170375; PMID:8124714
A:Accession: A53193
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-418 <R0E>
A:Cross-References: GB:L27585; NID:9452159; PID:9452160
C:Superfamily: sonic hedgehog protein

```

Query Match 66.3%; Score 1517; DB 2; Length 418;

Best Local Similarity 67.7%; Pred. No. 9, 6e-115;

Matches 294; Conservative 47; Mismatches 75; Indels 18; Gaps 3;

```

OY 4 LARCELVTLASSLVCPGLACGPRGFRGRHPRKRLTFLAYKQFIPTVAEKTLCASGRY 63
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3 LTRVLVSLTLTLVSLVSGIACGPRGFRGRHPRKRLTFLAYKQFIPTVAEKTLCASGRY 62
OY 64 EGRITNSRERFELTPNYPDIIFKDEENTGADRLMTORCKDKLNLAISVNMHPGVRL 123
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 63 EGRITNSRERFELTPNYPDIIFKDEENTGADRLMTORCKDKLNLAISVNMHPGVRL 122
OY 124 RTEGDEDEGHHSESLHYEGRAVDITTSDRDSKYGMRLAVEAGPDMVYVESKAH 183
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 123 RTEGDEDEGHHSESLHYEGRAVDITTSDRDSKYGMRLAVEAGPDMVYVESKAH 182
OY 184 CSYKAENSVAAKSGCGFPGSATVHLEBOGGTKLVKDLRPGDRVLAADOGRLTSDFL 243
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 183 CSYKAENSVAAKSGCGFPGSALVSLDGGOKAVKDLNPDQVLAADSAGNLVFSDDIMPT 242
OY 244 DRDEGAKKVFYIETLEPERELLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQRY 303
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 243 DRDSTTRRFVYIETQEPVEKITLTAHLLFVL-DNSTEDLHTMTAAYSSVRAQGVAV 301
OY 304 VAERGGDRRLPAAVHSVTLREBEAGAVAPLTAGHTILINRVLASCYAVIEEHSMAHRA 363
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 302 VDSGOLKSVIYORIY---EQRGSFAPVTAGHTIVVDRIASCYAVIEDGGLAHIAF 357
OY 364 AFPRLAHLAALAPARTDGGGGGSIPTAAOSATEARGAEPTAGIHMYSQLYHIGTWL 423
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 358 AARLLYVSSFLP-----QNSSRSNATLQOEIVHYSRLTLQMGTWL 404
OY 424 SETMHPGLMAVSS 437
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

DB 405 SNMLHPGLMSVSS 418

RESULT 6

A49426

sonic hedgehog gene shh protein - zebra fish

N:Alternate names: local cell-cell interaction signaling protein

C:Species: Brachydanio rerio (zebra fish)

C>Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 10-Dec-1999

C:Accession: A49426

R:Krauss, S.; Concordet, J.P.; Ingham, P.W.

A:Title: A functionally conserved homolog of the Drosophila segment polarity gene hh

A:Reference number: A49426; MUID:94094335; PMID:826519

A:Accession: A49426

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-415 <KRA>

A:Experimental source: embryo

A>Note: sequence extracted from NCBI backbone (NCBIP:142459)

C:Superfamily: sonic hedgehog protein

Query Match 63.7%; Score 1458; DB 2; Length 415;

Best Local Similarity 67.1%; Pred. No. 5, 6e-110;

Matches 289; Conservative 46; Mismatches 78; Indels 18; Gaps 5;

```

OY 4 LARCELVTLASSLVCPGLACGPRGFRGRHPRKRLTFLAYKQFIPTVAEKTLCASGRY 63
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3 LTRVLVSLTLTLVSLVSGIACGPRGFRGRHPRKRLTFLAYKQFIPTVAEKTLCASGRY 62
OY 64 EGRITNSRERFELTPNYPDIIFKDEENTGADRLMTORCKDKLNLAISVNMHPGVRL 123
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 63 EGRITNSRERFELTPNYPDIIFKDEENTGADRLMTORCKDKLNLAISVNMHPGVRL 122
OY 124 RTEGDEDEGHHSESLHYEGRAVDITTSDRDSKYGMRLAVEAGPDMVYVESKAH 183
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 123 RTEGDEDEGHHSESLHYEGRAVDITTSDRDSKYGMRLAVEAGPDMVYVESKAH 182
OY 184 CSYKAENSVAAKSGCGFPGSATVHLEBOGGTKLVKDLRPGDRVLAADOGRLTSDFL 243
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 183 CSYKAENSVAAKSGCGFPGSALVSLDGGOKAVKDLNPDQVLAADSAGNLVFSDDIMPT 242
OY 244 DRDEGAKKVFYIETLEPERELLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQRY 303
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 243 DRDSTTRRFVYIETQEPVEKITLTAHLLFVL-DNSTEDLHTMTAAYSSVRAQGVAV 301
OY 304 VAERGGDRRLPAAVHSVTLREBEAGAVAPLTAGHTILINRVLASCYAVIEEHSMAHRA 363
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 302 VDSGOLKSVIYORIY---EQRGSFAPVTAGHTIVVDRIASCYAVIEDGGLAHIAF 357
OY 364 AFPRLAHLAALAPARTDGGGGGSIPTAAOSATEARGAEPTAGIHMYSQLYHIGTWL 423
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 358 AARLLYVSSFLP-----PMRLYNRGSGTPTG-----SCHQMGTW 404
OY 424 SETMHPGLMAV 434
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 405 SNMLHPGLMSV 415

```

RESULT 7

B49425

Desert hedgehog protein precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999

R:Reichardt, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; M

Cell 75, 1417-1430, 1993

A:Title: Sonic hedgehog, a member of a family of putative signaling molecules, is imp

A:Reference number: B49425; MUID:94094344; PMID:7916661

A:Accession: B49425

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-396 <ECH>

A:Cross-references: GB:X76292; NID:g443941; PIDN:CAA53924.1; PID:g443942
 C:Genetics:
 A:Gene: Dhh
 C:Superfamily: sonic hedgehog protein

Query Match 53.3%; Score 1219; DB 2; Length 396;
 Best Local Similarity 58.7%; Pred. No. 11e-90;
 Matches 249; Conservative 55; Mismatches 80; Indels 40; Gaps 12;

QY 4 LLARCFVLIASSLLVCGIACGPGRG-FGKRHRPK-LTPYAKOTIPNVAEKTIGASG 61
 DB LPLCLALALAS-----AOSGPGRGVRRRYRKOLVPLTYQFSPMERLIGASG 61
 QY 7 REGKITRNSERFKELTPNYNDITFKDEBNTGADRLMTORCKDLNALAISVNMQPGV 121
 DB 62 PAEGVTRKSERFRLVNPNDITFKDEBNSGADRLMTERCKERYNALAISVNMMPGV 121
 QY 122 RLRTVEGDEDEHSEESLHYEGRAVDITTSDRDRSKYGMRLAVEAGFDWYVESKAH 181
 DB 122 RLRTVEGDEDEHMODSLHTEGRALDITTSDRDRNKITGLRLAVEAGFDWYVESRHH 181
 QY 182 IHCYKAENSVAAKSGCGFPGSATVHLBOGCTKLVKDLRPGDRLVLAADQGRILYSDEL 241
 DB 182 IHVYKADNSLAVRAGCGFPGMATVRLRSGERKGLRELHSGDWLAADAAGVPTPVLL 241
 QY 242 FLDRBEGAKYFYVETLEPRERLLITAAHLIFVAPHNDSEPTGP-----SALFASRVBP 298
 DB 242 FLDRLOQRASFAVETERPPRKLLITPMHLYFAA-----RCGPAPADGFAPVFAHRLRAG 297
 QY 299 QRVYVVARSGRRLLPAVHSVTLREBAGAVAPLTAHGTILIRVLAASCAVIEESHW 358
 DB 298 DSVLA---PGGD-ALQPARVAVR-REBAGVFAPLTAHGTILVNDVLAASCAVIESHOW 352
 QY 359 AHRAAPRLAHLALAPARTDGGCGSIPAAQSATEARGAEPPTAGIHWYSQLYHIG 418
 DB 353 AHRAAPRLAHLA-LGALLP-----GGAV-----QPT-GMHWSRLLYALA 391
 QY 419 TWLL 422
 DB 392 EELM 395

RESULT 8

C49425
 Indian hedgehog protein - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 11-Jan-2000
 C:Accession: C49425
 R:Exchard, Y.; Egesten, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; McMahon, J.A.; Title: Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the development of the mouse embryo.
 A:Reference number: A49425; MUID:94094334; PMID:7916661
 A:Accession: C49425
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-336 <ECH>
 A:Cross-references: GB:X76291
 C:Superfamily: sonic hedgehog protein

Query Match 45.6%; Score 1043; DB 2; Length 336;
 Best Local Similarity 58.7%; Pred. No. 1.5e-76;
 Matches 213; Conservative 36; Mismatches 80; Indels 32; Gaps 5;

QY 72 ERFKELPNYNDITFKDEBNTGADRLMTORCKDLNALAISVNMQPGVRLRYTEGDE 131
 DB 1 ERFKELPNYNDITFKDEBNTGADRLMTORCKDLNALAISVNMQPGVRLRYTEGDE 60
 QY 132 DGHSEESLHAGRAVDITTSDRDRSKYGMRLAVEAGFDWYVESKAHITCSVKAENS 191
 DB 61 DGHSEESLHAGRAVDITTSDRDRNKYGLRLAVEAGFDWYVESKAHVCYSKSHS 120
 QY 192 VAAKSGCFPGSATVHLBOGCTKLVKDLRPGDRLVLAADQGRILYSDELTLDRDEGAKK 251
 DB 192 VAAKSGCFPGSATVHLBOGCTKLVKDLRPGDRLVLAADQGRILYSDELTLDRDEGAKK 251

DB 121 AAKTGCGFPAQAQVRLNENGERVALSAVKPGDRLVLAAGEDTPTFSVLIPLDRPNRLR 180
 QY 252 VEVYIETLEPRERLLITAAHLIFVAPHNDSEPTGPSPALFASRVPGGVVVAERGGDR 311
 DB 181 AFOVETQDPPRRLLITAAHLIFVAPHNDSEPTGPSPALFASRVPGGVVVAERGGDR 236
 QY 312 RLPAVHVSVTLREBAGAVAPLTAHGTILIRVLAASCAVIEESHWRAFAFRLAHA 371
 DB 237 LQPARVAAS-THVALGSYAPLRTHTLVVEDVVASCFAAVADHNLQLAFWPLRLRPS 294
 QY 372 LLAALAPARTDGGCGSIPAAQSATEARGA-EPTAGIHWYSQLYHIGTILYDSTHPL 430
 DB 295 L-----AWGSWTPSEGVHSTPOMLYRLGLRLLLESTFHL 329
 QY 431 GMA 433
 DB 330 GMS 332

RESULT 9

A46400
 segment polarity protein hedgehog - fruit fly (Drosophila melanogaster)

N:Alternate names: hh protein
 C:Species: Drosophila melanogaster
 C:Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 10-Dec-1999
 C:Accession: A46400; JN0501; A43480
 R:Tabata, T.; Eaton, S.; Kornberg, T.B.
 Gene Dev. 6, 2635-2645, 1992

A:Title: The Drosophila hedgehog gene is expressed specifically in posterior compartment
 A:Reference number: A46400; MUID:94040725; PMID:1340474
 A:Accession: A46400
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-471 <TAB>
 A:Cross-references: GB:566384; NID:g435848; PID:g435849
 A:Experimental source: Oregon R
 R:Tashiro, S.; Michie, T.; Hiyashijima, S.; Zennaro, S.; Ishimaru, S.; Takahashi, F.; Gene 124, 183-189, 1993
 A:Title: Structure and expression of hedgehog, a Drosophila segment-polarity gene req
 A:Reference number: JN0501; MUID:93185922; PMID:8166882
 A:Accession: JN0501

A:Molecule type: mRNA
 A:Residues: 1-471 <TAB>
 A:Cross-references: GB:I05404
 A:Note: it is uncertain whether Met-1 or Met-51 is the initiator
 R:Lee, J.J.; von Kessler, D.P.; Parks, S.; Beachy, P.A.
 Cell 71, 33-50, 1992
 A:Title: Secretion and localized transcription suggest a role in positional signaling
 A:Reference number: A43480; MUID:93008241; PMID:13944530
 A:Accession: A43480
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-471 <LEB>
 A:Cross-references: GB:L02793; NID:g157609; PID:g157610
 A:Note: sequence extracted from NCBI backbone (NCBI:P:115418)
 C:Comment: This protein is required for cell-cell communication.
 C:Genetics:
 A:Gene: hh
 A:Cross-references: Flybase:FBgn0004644
 A:Initons: 160/3; 248/1
 C:Superfamily: sonic hedgehog protein
 C:Keywords: transmembrane protein
 F:62-82/Domain: transmembrane #status predicted <TKM>

Query Match 39.1%; Score 894; DB 2; Length 471;
 Best Local Similarity 46.8%; Pred. No. 2.5e-64;
 Matches 197; Conservative 62; Mismatches 140; Indels 22; Gaps 9;

QY 10 LVIIASSLVPGIACGPGRGFGKRHRPKLTPYAKOTIPNVAEKTIGASGREGKITR 69
 DB 70 LLLIVLPVESPASGPGRGGLRRH-ARNLVPLVLTQITPNSLEYTNSASGPLEGVIRR 128

Db 1116 SHLESBS-AHKLIIDF 1130

RESULT 13

724045

hypothetical protein R08B4.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T24045

R:White, S.

submitted to the EMBL Data Library, November 1995

A:Reference number: Z19834

A:Accession: T24045

A:Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1226 <WILL>

A:Cross-references: EMBL:Z68008; PIDN:CA92000.1; GSPDB:GN00028; CESP:R08B4.1

A:Experimental source: clone R08B4

C:Genetics:

A:Gene: CESP:R08B4.1

A:Map position: X

A:Introns: 53/2; 113/1; 152/3; 204/3; 272/1; 354/1; 389/3; 714/3; 839/3; 877/1; 960/3; 1

Query Match

Best Local Similarity 26.8%; Score 148; DB 2; Length 1226;

Matches 51; Conservative 39; Mismatches 66; Indels 34; Gaps 9;

Db 190 NSVAKSGCGPGSATVHLGGTKLVKDLRPGDRVLAADDOGRLLYSDFLFLDRDEGA 249

Db 1012 SALVATACGCSLDTWV-TTPGKKRMOQIDIGYVLADE-KYFTPTILMIREPEK 1069

QY 250 KKFVYVLEPRERLLTAHLF-----VAPHNSGPTPGPSALFASRYR 296

Db 1070 VQEFLLTMT-EXGKTLRTSRHFMYRNKCGKSPQYIKMLPHDG-----EAFASDLE 1121

QY 297 PGQRYVAERGGDRLLPAVHSTLREBGAAPLTAHGTILINRYLASCYVIEEH 356

Db 1122 VDDCV-VVLYKGYKQ---QKLETT-RSVRTGYSPLTNNGRITVNDMLASCYSEIQON 1176

QY 357 S-----MAH 360

Db 1177 TLQTFEWMAY 1186

RESULT 14

727665

hypothetical protein ZK1037.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27665

R:Baslam, V.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z20401

A:Accession: T27665

A:Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-481 <WILL>

A:Cross-references: EMBL:Z81142; PIDN:CA803509.1; GSPDB:GN00023; CESP:ZK1037.10

A:Experimental source: clone ZK1037

C:Genetics:

A:Gene: CESP:ZK1037.10

A:Map position: 5

A:Introns: 37/3; 75/3; 141/3; 173/2; 253/1; 298/3; 310/1; 371/2; 424/3

Query Match

Best Local Similarity 25.6%; Score 146; DB 2; Length 481;

Matches 43; Conservative 35; Mismatches 70; Indels 20; Gaps 5;

QY 199 CFPGSATVHLGGTKLVKDLRPGDRVLAADDOGRLLYSDFLFL-----DRDEGAK 250

Db 314 CFPNDAYVNVYEKAVKRDELEIGDWVADENG-----EDITFLPYKYLHHRDEQENAE 368

QY 251 KVFVYVLEPRERLLTAHLFVAP--HNSGPTPGPSALFASRYRPGQRYVVAERG 308

Db 369 ---FILESLDNGEFTLTKELHLYTTECRONSSSELKISMSISACKVAGCOFYLAQSEA 425

QY 309 GDRLLPAVHSTLREBGAAPLTAHGTILINRYLASCYVIEEH 356

Db 426 LTKRYLVEILD--IKRVKGTGYAPMTSQGHLLVKNKIHTSCHSEVDHH 471

RESULT 15

719563

hypothetical protein C29F3.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999

C:Accession: T19563

R:Matthews, L.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19142

A:Accession: T19563

A:Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-629 <WILL>

A:Cross-references: EMBL:Z81043; PIDN:CA802804.1; GSPDB:GN00023; CESP:C29F3.2

A:Experimental source: clone C29F3

R:White, S.

submitted to the EMBL Data Library, June 1998

A:Reference number: Z19657

A:Accession: T23034

A:Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-629 <WILL>

A:Cross-references: EMBL:AL023813; PIDN:CA19424.1; GSPDB:GN00023; CESP:C29F3.2

A:Experimental source: clone H02K04

C:Genetics:

A:Gene: CESP:C29F3.2

A:Map position: 5

A:Introns: 23/1; 111/3; 177/3; 207/2; 287/1; 381/3; 399/3; 417/1; 476/2; 528/3; 537/2

Query Match

Best Local Similarity 27.3%; Score 140.5; DB 2; Length 629;

Matches 45; Conservative 31; Mismatches 72; Indels 17; Gaps 6;

QY 199 CFPGSATVHLGGTKLVKDLRPGDRVLAADDOGRLLYSDFLFLDRDEGAKKVFYIET 258

Db 421 CFPADAEVNVYEKAVKRDELEVGDMVQALHGK-ETTVSPVKYWLHRDP-EQEAFFVEFL 478

QY 259 LEPRERLLTAHLF-----VAPHNSGPTPGPSALFASRYRPGQRYVVAERGGDR 311

Db 479 LENGESFTLTKHLVFATDCQONKKNLDLNPST-----TKINIGE-CFMAQPENAS 531

QY 312 RLPAVHSTLREBGAAPLTAHGTILINRYLASCYVIEEH 356

Db 532 KFKQVQILDIG-RVRKGTGYAPMTSLGHLVNOIHITSCHSEIDHH 575

Search completed: February 20, 2003, 10:13:34
Job time : 18.3815 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:07:32 ; Search time 13.4009 seconds
(without alignments)
1352.533 Million cell updates/sec

Title: US-09-827-110A-13

Perfect score: 2289
Sequence: 1 MLLLRARCPVLIIASSLLVC.....GTWLDSFTMHPLGNVAKSS 437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2286	99.9	437	1 SHH_MOUSE	062226 mus musculus
2	2200	96.1	437	1 SHH_RAT	063673 rattus norv
3	2065	90.2	462	1 SHH_HUMAN	013465 homo sapien
4	1788	78.1	425	1 SHH_CHICK	091035 gallus gall
5	1542.5	67.4	432	1 SHH_XENLA	090385 cynops pyr
6	1525.5	66.6	444	1 SHH_XENLA	092000 xenopus lae
7	1517	66.3	418	1 SHH_BRARE	092008 brachydantio
8	1464	64.0	416	1 SHH_BRARE	090419 brachydantio
9	1310	57.2	408	1 IHH_CHICK	098938 gallus gall
10	1271.5	55.5	411	1 IHH_HUMAN	014623 homo sapien
11	1268	55.4	411	1 IHH_MOUSE	097812 mus musculu
12	1267	55.4	409	1 IHH_XENLA	091632 xenopus lae
13	1228	53.6	412	1 IHH_BRARE	098862 brachydantio
14	1219	53.3	386	1 DHH_MOUSE	061488 mus musculu
15	1217	53.2	386	1 DHH_HUMAN	043333 homo sapien
16	1104.5	48.3	398	1 DHH2_XENLA	091611 xenopus lae
17	1098.5	48.0	396	1 DHH1_XENLA	002936 drosophila
18	894	39.1	471	1 HH_DROME	002936 drosophila
19	855.5	37.4	481	1 HH_DROME	002936 drosophila
20	603	26.3	421	1 SHH_RASEL	079868 raspora ele
21	600	26.2	421	1 SHH_CARAU	079691 carassius a
22	600	26.2	421	1 SHH_PUNTE	079850 puntius tet
23	599	26.2	421	1 SHH_RASHE	079864 raspora het
24	599	26.2	421	1 SHH_RASPA	079869 raspora pav
25	598	26.1	421	1 SHH_TANAL	079915 tanichthys
26	596	26.0	421	1 SHH_AMCH	079662 amblypharyn
27	595	26.0	421	1 SHH_DANNA	013235 danio aff.
28	595	26.0	421	1 SHH_DANNA	013235 danio aff.
29	595	26.0	421	1 SHH_DANAT	013238 danio aff.
30	595	26.0	421	1 SHH_DANER	013245 danio frank
31	595	26.0	421	1 SHH_DANKE	079709 danio kerrl
32	595	26.0	421	1 SHH_DANPU	079717 danio pulch
33	595	26.0	421	1 SHH_DEYDE	013241 devario dev

34	595	26.0	121	1 SHH_DEYMA	013247 devario mal
35	595	26.0	121	1 SHH_DEYPA	013250 devario pat
36	591	25.8	121	1 SHH_PUNCO	079838 puntius con
37	399	17.4	88	1 DHH_BRARE	079729 brachydantio
38	286	12.9	58	1 IHH_CARAU	079693 carassius a
39	286	12.9	58	1 IHH_DANAT	013240 danio aff.
40	296	12.9	58	1 IHH_DANKE	079719 danio kerrl
41	296	12.9	58	1 IHH_DANPU	079719 danio pulch
42	296	12.9	58	1 IHH_DEYDE	013243 devario dev
43	296	12.9	58	1 IHH_PUNTE	079852 puntius tet
44	286	12.9	58	1 IHH_RASEL	079860 raspora ele
45	291	12.7	58	1 IHH_DEYMA	013215 devario ma1

ALIGNMENTS

RESULT 1
SHH_MOUSE
ID 062226;
AC 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sonic hedgehog protein precursor (SHH) (HHG-1).
GN SHH OR HHG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=94094334; PubMed=7916661;
RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Moller J.,
RA McMahon J.A., McMahon A.P.,
RT "Sonic hedgehog, a member of a family of putative signaling
molecules, is implicated in the regulation of CNS polarity.";
RL Cell 75:1417-1430(1993).
RN [2]
RP REVISION TO 122.
RC STRAIN=C57BL/6J;
RA McMahon A.P.,
RN Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RX [3]
RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
RX MEDLINE=95236997; PubMed=7720571;
RA Zhang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
RA Zhao R., Seidlin M.F., Fallon J.F., Beachy P.A.,
RT "Products, genetic linkage and limb patterning activity of a murine
hedgehog gene.";
RL Development 120:3339-3353(1994).
RN [4]
RP FUNCTION, AND AUTOPROTEOLYTIC CLEAVAGE.
RX MEDLINE=95254654; PubMed=7736596;
RA Roelink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,
RA Beachy P.A., Jessell T.M.,
RT "Floor plate and motor neuron induction by different concentrations of
the amino-terminal cleavage product of sonic hedgehog
autoproteolysis.";
RL Cell 81:445-455(1995).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 34-195.
RX MEDLINE=96069744; PubMed=7477329;
RA Hall T.M.T., Porter J.A., Beachy P.A., Leahy D.J.,
RT "A potential catalytic site revealed by the 1.7-A crystal structure of
the amino-terminal signalling domain of Sonic hedgehog.";
RL Nature 378:212-216(1995).
CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION
OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESENTS THE
CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
TARGET, THE G11 ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A

VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY SIMILARITY).

-1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).

-1- TISSUE SPECIFICITY: EXPRESSED IN A NUMBER OF EMBRYONIC TISSUES INCLUDING THE NOTOCHORD, VENTRAL NEURAL TUBE, FLOOR PLATE, LUNG BUD, ZONE OF POLARIZING ACTIVITY AND POSTERIOR DISTAL MESENCHYME OF LIMBS. IN THE ADULT, EXPRESSED IN LUNG AND NEURAL RETINA.

-1- DEVELOPMENTAL STAGE: FIRST DETECTABLE DURING GASTRULATION.

-1- INDUCTION: BY RETINOIC ACID.

-1- PM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY.

-1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.

THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION - THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL ENTITIES REQUIRES A LICENSE AGREEMENT (SEE <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: X76290; CAA53922.1; -
 PDB: 1VHH; 29-JAN-96.
 DR MEROPS; C46.002; -
 DR MGD; MG1.98297; Shh.
 DR InterPro: IPR000320; HH_signal.
 DR InterPro: IPR001767; Hedgehog_hnt.
 DR InterPro: IPR003586; Hedgehog_hntc.
 DR InterPro: IPR003587; Hedgehog_hntn.
 DR InterPro: IPR002203; Intein.
 DR InterPro: IPR001657; SonicHH.
 DR Pfam: PF01079; Hnt; 1.
 DR Pfam: PF01085; HH_signal; 1.
 DR PRINTS; PR00632; SONICHHOG.
 DR ProDom; PD003042; HH_signal; 1.
 DR SMART; SM00305; Hntc; 1.
 DR SMART; SM00306; Hntn; 1.
 DR PROSITE; PS50817; INTEIN_NTER; 1.
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal; Lipoprotein; Palmitate; 3d-structure.
 FT SIGNAL 1 24
 FT CHAIN 25 437
 FT CHAIN 25 198
 FT CHAIN 199 437
 FT SITE 198 199
 FT SITE 244 244
 FT SITE 268 268
 FT ACT_SITE 271 271
 FT BINDING 198 198
 FT DOMAIN 383 387
 FT LIPID 25 25
 FT SEQUENCE 437 AA; 47773 MW; DOB872F08E7860EE CRC64;

Query Match 99.9%; Score 2286; DB 1; Length 437;

Best Local Similarity 99.9%; Pred. No. 3e-176;
 Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MILLARCEFLVILASSLLVCGIACGPGRGKRRHPKPLIPYAKOTIPVAVETKIGAS 60
 1 MILLARCEFLVILASSLLVCGIACGPGRGKRRHPKPLIPYAKOTIPVAVETKIGAS 60

Db 1 MILLARCEFLVILASSLLVCGIACGPGRGKRRHPKPLIPYAKOTIPVAVETKIGAS 60

Oy 61 GRYEGRITRNSERFELTPNPNPDIFRDEENTGADRLMTORCKDKLNAISVNWMPG 120
 61 GRYEGRITRNSERFELTPNPNPDIFRDEENTGADRLMTORCKDKLNAISVNWMPG 120

Db 61 GRYEGRITRNSERFELTPNPNPDIFRDEENTGADRLMTORCKDKLNAISVNWMPG 120

Oy 121 VRLRTGEMDEDGHHSESLIYEGAVDITTSDBRSRYGMALRLAVAGDWTYTESKA 180
 121 VRLRTGEMDEDGHHSESLIYEGAVDITTSDBRSRYGMALRLAVAGDWTYTESKA 180

Db 121 VRLRTGEMDEDGHHSESLIYEGAVDITTSDBRSRYGMALRLAVAGDWTYTESKA 180

Oy 181 HHCVRKENSVAAKSGCGPGSATVHLDEGGTKLVKLRGDRVLADDDGRLIYSFL 240
 181 HHCVRKENSVAAKSGCGPGSATVHLDEGGTKLVKLRGDRVLADDDGRLIYSFL 240

Db 181 HHCVRKENSVAAKSGCGPGSATVHLDEGGTKLVKLRGDRVLADDDGRLIYSFL 240

Oy 241 TFLDRDEGAKRVFYIETLEPRERLLTAHLFVAPHNDSGPTPGSPALFASRYRPGQR 300
 241 TFLDRDEGAKRVFYIETLEPRERLLTAHLFVAPHNDSGPTPGSPALFASRYRPGQR 300

Db 241 TFLDRDEGAKRVFYIETLEPRERLLTAHLFVAPHNDSGPTPGSPALFASRYRPGQR 300

Oy 301 VYVAERGGDRRLPAVAVSTLREEGATAPLTAHTIILNRYLASCYAVIEBSMAH 360
 301 VYVAERGGDRRLPAVAVSTLREEGATAPLTAHTIILNRYLASCYAVIEBSMAH 360

Db 301 VYVAERGGDRRLPAVAVSTLREEGATAPLTAHTIILNRYLASCYAVIEBSMAH 360

Oy 361 RAFAFRILAHMLALAPARTDGGGSIIPAOSATEARGAEPTAGIMYSOLLHYHGTW 420
 361 RAFAFRILAHMLALAPARTDGGGSIIPAOSATEARGAEPTAGIMYSOLLHYHGTW 420

Db 361 RAFAFRILAHMLALAPARTDGGGSIIPAOSATEARGAEPTAGIMYSOLLHYHGTW 420

Oy 421 LIDSETMHPILGNVAKSS 437
 421 LIDSETMHPILGNVAKSS 437

Db 421 LIDSETMHPILGNVAKSS 437

RESULT 2

SHH_RAT SHH_RAT STANDARD; PRT; 437 AA.
 AC 063673;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sonic hedgehog protein precursor (SHH).
 GN SHH OR VHH-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Embryonic floor plate;
 RX MEDLINE=94170375; PubMed=8124714;
 RA Roelink H., Augsburg A., Heemskerk J., Korzh V., Norlin S.,
 RA Ruiz I., Altaba A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,
 RA Dodd J.;
 RT "Floor plate and motor neuron induction by vhh-1, a vertebrate homolog
 of hedgehog expressed by the notochord.";
 RL Cell 76:761-775(1994).
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
 CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
 CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE

CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE NODE, NOTOCHORD, FLOOR PLATE,
 CC AND POSTERIOR LIMB BUD MESCENCHYME.
 CC -1- PFM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L27340; AAA20999.1; -
 CC HSSP: Q62226; 1VHH.
 CC DR MEROPS: C46.002; -
 CC DR InterPro: IPR0000320; HH_signal.
 CC DR InterPro: IPR001767; Hedgehog_hnt.
 CC DR InterPro: IPR003586; Hedgehog_hntc.
 CC DR InterPro: IPR003587; Hedgehog_hlntn.
 CC DR InterPro: IPR002203; Inteln.
 CC DR InterPro: IPR001657; SonicHH.
 CC DR Pfam: PF01079; Hnt: 1.
 CC DR Pfam: PF01085; HH_signal: 1.
 CC DR PRINTS: PR00632; SONICHHOG.
 CC DR PRODOM: PD003042; HH_signal: 1.
 CC DR SMART: SM00305; Hntc: 1.
 CC DR SMART: SM00306; Hntcn: 1.
 CC DR PROSITE: PS0817; INTERIN_N_TER: 1.
 CC KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 CC Signal; Lipoprotein; Palmitate.
 CC KM SIGNAL 1 24
 CC FT CHAIN 25 437 SONIC HEDGEHOG PROTEIN.
 CC FT CHAIN 25 198 SONIC HEDGEHOG PROTEIN N-PRODUCT.
 CC FT CHAIN 199 437 SONIC HEDGEHOG PROTEIN C-PRODUCT.
 CC FT SITE 198 199 CLEAVAGE (AUTO-).
 CC FT SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY
 CC FT SITE SIMILARITY).
 CC FT SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY
 CC FT SITE SIMILARITY).
 CC FT ACT_SITE 271 271 ESSENTIAL FOR AUTO-CLEAVAGE (BY
 CC FT SITE SIMILARITY).
 CC FT BINDING 198 198 CHOLESTEROL (BY SIMILARITY).
 CC FT DOMAIN 383 387 POLY-GLY.
 CC FT LIPID 25 25 PALMITATE (BY SIMILARITY).
 CC SQ SEQUENCE 437 AA; 47630 MW; 0DBFC19F0D1662A0 CRC64;
 CC
 CC Query Match 96.1%; Score 2200; DB 1; Length 437;
 CC Best Local Similarity 96.3%; Pred. No. 2,4e-16; Mismatches 13; Indels 0; Gaps 0;
 CC Matches 421; Conservative 3;
 CC DB 1 MILLARCPFLVILASSLAVCPGLACPGRGFGKRRHPKPLPLAVKOPFIPNVAEKTIGS 60
 CC DB 1 MILLARCPFLVILASSLAVCPGLACPGRGFGKRRHPKPLPLAVKOPFIPNVAEKTIGS 60
 CC DB 61 GYEGKITRNSRFKELIPNVPDIIIFKDENTGADRLMTQRCOKDLNALAISVNMQMPG 120
 CC DB 61 GYEGKITRNSRFKELIPNVPDIIIFKDENTGADRLMTQRCOKDLNALAISVNMQMPG 120
 CC DB 61 GYEGKITRNSRFKELIPNVPDIIIFKDENTGADRLMTQRCOKDLNALAISVNMQMPG 120
 CC DB 121 VALRTGEGDEGHHSESLHYEGRAVDITTSDRDSKYGMLARLAVEGPMVYTESKA 180
 CC DB 121 VALRTGEGDEGHHSESLHYEGRAVDITTSDRDSKYGMLARLAVEGPMVYTESKA 180

QY 181 HIKSVKAENSVAAKSGCGPFSATVHLEGGTKLVKDLRPGDRVLAADDOGRLLYSDFL 240
 DB 181 RHICSVKAENSVAAKSGCGPFSATVHLEGGTKLVKDLRPGDRVLAADDOGRLLYSDFL 240
 QY 241 TFLDBESAKKVFYIEFLPERELLTAHLLEVPAPINDSGPPPGSPALFASRVPCOR 300
 DB 241 TFLDBESAKKVFYIEFLPERELLTAHLLEVPAPINDSGPPPGSPALFASRVPCOR 300
 QY 301 VYVAERGGDRRLPAVAHSTVLRREAGAVAPLTAHGTILINRVLASCVAIEEHSMAH 360
 DB 301 VYVAERGGDRRLPAVAHSTVLRREAGAVAPLTAHGTILINRVLASCVAIEEHSMAH 360
 QY 361 RAFAFRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPPTAGIMYSOLYHIGTW 420
 DB 361 RAFAFRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPPTAGIMYSOLYHIGTW 420
 QY 421 LIDSETMHPGLMAVKS 437
 DB 421 LIDSETMHPGLMAVKS 437
 RESULT 3
 ID SHH_HUMAN STANDARD; PRT; 462 AA.
 AC Q15465;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sonic hedgehog protein precursor (SHH) (HhG-1).
 GN SHH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal lung;
 RX MEDLINE=96070431; PubMed=7590746;
 RA Marigo V., Roberts D.J., Lee S.M.K., Tsukurov O., Leyl T.,
 RA Gaster J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seldman C.E.,
 RA Jenkins N.A., Seidman J.G., McMahon A.P., Tabin C.;
 RT "Cloning, expression, and chromosomal location of SHH and HH: two
 RT human homologues of the Drosophila segment polarity gene hedgehog.";
 RN Genomics 28:44-51(1995).
 [2]
 RP SEQUENCE FROM N.A.
 RA Tate G., Kishimoto K., Mitsuya T.;
 RT "Expression of Sonic hedgehog and its receptor Patched/Smoothed in
 RT human cancer cell lines and embryonic organs.";
 RN J. Biochem. Mol. Biol. Biophys. 4:27-34(2000).
 [3]
 RP SEQUENCE OF 1-187 FROM N.A.
 RA Strong C., Graves T., Sutterer C., Ozeraky P.;
 RL Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.
 [4]
 RP SEQUENCE OF 119-167 FROM N.A.
 RX MEDLINE=95236997; PubMed=7720571;
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
 RA Zhao R., Seidman M.F., Fallon J.F., Beachy P.A.;
 RT "Products, genetic linkage and limb patterning activity of a murine
 RT hedgehog gene.";
 RN Development 120:3339-3353(1994).
 [5]
 RP PALMITOYLATION.
 RX MEDLINE=98256339; PubMed=9593755;
 RA Pepinsky R.B., Zeng C., Wen D., Rayhorn P., Baker D.P., Williams K.P.,
 RA Bixler S.A., Ambrose C.M., Gardner E.A., Miatkowski K., Taylor F.R.,
 RA Wang E.A., Galdes A.;
 RT "Identification of a palmitic acid-modified form of human Sonic
 RT hedgehog.";
 RN J. Biol. Chem. 273:14037-14045(1998).
 [6]

RP VARIANTS HPE ARG-31; GLY-117; ARG-117.
RX MEDLINE-97051937; PubMed-8896572;
RA Roessler E., Belloni E., Gaudenz K., Jay P., Berta P., Scherer S.W.,
Tsui L.-C., Muenke M.;
RT "Mutations in the human Sonic hedgehog gene cause holoprosencephaly.";
RL Nat. Genet. 14:357-360(1996).
RN [7].
RP VARIANTS HPE ARG-31; GLY-117; ARG-117; GLU-224; THR-226 AND THR-383.
RX MEDLINE-98027056; PubMed-9302822;
RA Roessler E., Belloni E., Gaudenz K., Vargas F., Scherer S.W.,
Tsui L.-C., Muenke M.;
RT "Mutations in the C-terminal domain of Sonic hedgehog cause holoprosencephaly.";
RL Hum. Mol. Genet. 6:1847-1853(1997).
RN [8].
RP VARIANTS HPE HIS-100; GLN-188 AND ASN-222.
RX MEDLINE-99371775; PubMed-10441331;
RA Odent S., Acti-Bitach T., Blayau M., Mathieu M., Aug J.,
Delezo de A.L., Gall J.Y., Le Marcq B., Munnich A., David V.,
Vekemans M.;
RT "Expression of the Sonic hedgehog (SHH) gene during early human development and phenotypic expression of new mutations causing holoprosencephaly.";
RL Hum. Mol. Genet. 8:1683-1689(1999).
RN [9].
RP VARIANTS HPE V-88; K-115; R-236; 263-R--A-269 DEL; D-290; A-424 AND L-436.
RX MEDLINE-20025757; PubMed-10556296;
RA Menall L., Ming J.E., Boclan M., Steinhans K., Bianchi D.W.,
Die-Smulders C., Giannotti A., Imtiaz M., Jones K.L., Campo M.D.,
Martin R.A., Melnick P., Pierpont M.E.M., Robin N.H., Young I.D.,
Roessler E., Muenke M.;
RT "The mutational spectrum of the sonic hedgehog gene in holoprosencephaly: SHH mutations cause a significant proportion of autosomal dominant holoprosencephaly.";
RL Hum. Mol. Genet. 8:2479-2488(1999).
RN [10].
RP FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BOD. DISPLAYS BOTH FLOOR PLATE AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY SIMILARITY).
CC SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC TISSUE SPECIFICITY: EXPRESSED IN FETAL INTESTINE, LIVER, LUNG, AND KIDNEY. NOT EXPRESSED IN ADULT TISSUES.
CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -1- DISEASE: DEFECTS IN SHH ARE THE CAUSE OF HOLOPROSENCEPHALY (HPE). HPE IS A CLINICALLY VARIABLE AND GENETICALLY HETEROGENEOUS MALFORMATION IN WHICH THE DEVELOPING FOREBRAIN FAILS TO CORRECTLY SEPARATE INTO RIGHT AND LEFT HEMISPHERES. IN THE MOST SEVERE FORM (ALOBA HPE), THERE IS NO INTERHEMISPHERIC FISSURE, A SINGLE BRAIN VENTRICLE IS PRESENT. HPE IS ASSOCIATED WITH SEVERAL DISTINCT FACIES AND PHENOTYPIC VARIABILITY. IN THE MOST EXTREME CASES, ANOPHTHALMIA OR CYCLOPIA IS EVIDENT ALONG WITH A CONGENITAL ABSENCE OF THE MATURE NOSE. THE LESS SEVERE FORM FEATURES FACIAL

	CC	DYSMORPHIA CHARACTERIZED BY OCULAR HYPERTELOMERISM, DEFECTS OF THE UPPER LIP AND/OR NOSE, AND ABSENCE OF THE OROFACIAL NERVES OR CORPUS CALLOSUM. THE MAJORITY OF HPE CASES ARE APPARENTLY SPORADIC, ALTHOUGH CLEAR EXAMPLES OF AUTOSOMAL DOMINANT (AD) INHERITANCE HAVE BEEN DESCRIBED. INTERESTINGLY, UP TO 30% OF OBLIGATE CARRIERS OF HPE GENE IN AD PEDIGREES ARE CLINICALLY UNAFFECTED.
	CC	-1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
	CC	-1- DATABASE: NAME=Atlas Genet. CytoGenet., OncoCan., Hematoli.; WWW= http://www.infobiogen.fr/services/chromancer/Genes/SHHD78.htm .
	CC	-----
	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-stb.ch/announce/or_send_an_email_to_license@isb-stb.ch).
	CC	-----
	DR	EMBL; L38518; AAA62179.1; --.
	DR	EMBL; AC002484; AAB67604.1; --.
	DR	HSSP; Q62226; LVHH.
	DR	MEROPS; C46_002; --.
	DR	GeneW; HGNC:10848; SHH.
	DR	MTM; 600725; --.
	DR	MTM; 142945; --.
	DR	InterPro: IPR000320; HH_signal.
	DR	InterPro: IPR001767; Hedgehog_hntc.
	DR	InterPro: IPR003586; Hedgehog_hntc.
	DR	InterPro: IPR003587; Hedgehog_hntn.
	DR	InterPro: IPR002203; Inteln.
	DR	InterPro: IPR001657; SonicHH.
	DR	Pfam; PF01079; Hntc_1.
	DR	Pfam; PF01085; HH_signal; 1.
	DR	PRINTS; PR00632; SONICHOG.
	DR	Prodom; PD003042; HH_signal; 1.
	DR	SMART; SM00305; Hntnc; 1.
	DR	SMART; SM00306; Hntcn; 1.
	DR	PROSITE; PS50817; INTEIN_N_TER; 1.
KM	Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;	
KM	Signal; Lipoprotein; Palmitate; Disease mutation; Holoprosencephaly.	
FT	SIGNAL	1 23
FT	CHAIN	24 462
FT	CHAIN	24 197
FT	CHAIN	198 462
FT	CHAIN	197 198
FT	SITE	243 243
FT	SITE	243 243
FT	SITE	267 267
FT	SITE	267 267
FT	ACT_SITE	270 270
FT	BINDING	197 197
FT	DOMAIN	407 411
FT	LIPID	24 24
FT	VARIANT	31 31
FT	VARIANT	88 88
FT	VARIANT	100 100
FT	VARIANT	115 115
FT	VARIANT	117 117
FT	VARIANT	117 117
FT	VARIANT	188 188
FT	VARIANT	222 222
FT	VARIANT	224 224
FT	VARIANT	224 224

FT VARIANT 226 226 A -> T (IN HBE; FAMILIAL).
 FT VARIANT 236 236 S -> R (IN HBE; FAMILIAL).
 FT VARIANT 263 269 /FTID-009170.
 Query Match 90.28; Score 2065; DB 1; Length 462;
 Best Local Similarity 87.28; Pred. No. 1,9e-158;
 Matches 403; Conservative 9; Mismatches 24; Indels 26; Caps 3;

2 LLLARCELYLASSLVCPGLGPGRGFGKRRPKPLPLAYKOPFPVNAEKTLAGS 61
 1 MLLARCLLVYSSLVSCGLACGPGRGFGKRRPKPLPLAYKOPFPVNAEKTLAGS 60
 62 RYEEKITRNSERFELTPNPNPDIIFDEENTGADRLMTORCKDKLALAIYVNNQPGV 121
 61 RYEEKITRNSERFELTPNPNPDIIFDEENTGADRLMTORCKDKLALAIYVNNQPGV 120
 122 RLRTGEMDEGHHSESLHTEGRAVDITTSDRSKYGMILAVENAGEDWYTESKAH 181
 121 KLRTGEMDEGHHSESLHTEGRAVDITTSDRSKYGMILAVENAGEDWYTESKAH 180
 182 IHCVKAENSYAAKSGGCPGSAIVHLEGGTKLVKDLRPGDRVLADDDGRLYSDEL 241
 181 IHCVKAENSYAAKSGGCPGSAIVHLEGGTKLVKDLRPGDRVLADDDGRLYSDEL 240
 242 FLDRDEGAKKVFYIETLEPERELLTAHLLFVAPHNH-----SGPTPG 286
 241 FLDRDEGAKKVFYIETLEPERELLTAHLLFVAPHNHDSATGEPASGSGPTPG 300
 287 PSALFASRVPRGQVYVVAERGRRLPAVHVSVTLEEPAGAVAPLHAGTLLINRYL 346
 301 PRAFLASRVPRGQVYVVAERGRRLPAVHVSVTLEEPAGAVAPLHAGTLLINRYL 360
 347 ASCAVVEESHMAHAPFRLAHLAALAPARTD-----GGGGGSP-AAOSA 395
 361 ASCAVVEESHMAHAPFRLAHLAALAPARTDGGSGGGGGRVLAIPA 420
 396 TEARGAEPAGIMYSOLLYHIGTWLDSMTMRPLGMAVKS 437
 421 ADARGAGATAGIMYSOLLYHIGTWLDSMTMRPLGMAVKS 462

RESULT 4
 ID SHH_CHICK STANDARD; PRT; 425 AA.
 AC 091035;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sonic hedgehog protein precursor (SHH).
 GN SHH.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 ON NCBI_TaxID=9031;
 RX MEDLINE=9409433; PubMed=8269518;
 RA Riddle R.D., Johnson R.L., Laufer E., Tabin C.;
 RT "Sonic hedgehog mediates the polarizing activity of the ZPA.";
 RL Cell 75:1401-1416(1993).
 RP FUNCTION. AND AUTOPROTEOLYTIC CLEAVAGE.
 RA MEDLINE=95254654; PubMed=7735596;
 RA Koellink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,
 RA Beachy P.A., Jessell T.M.;
 RT "Floor plate and motor neuron induction by different concentrations of
 RT the amino-terminal cleavage product of sonic hedgehog
 RT autoproteolysis.";
 RL Cell 81:445-455(1995).

CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
 CC ASSOCIATION WITH SMOOTHED (SMO), TO ACTIVATE THE TRANSCRIPTION
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
 CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE POSTERIOR LIMB BUD
 CC MESENCHYME, THE HENSEN'S NODE, THE NOTOCHORD, AND THE FLOOR PLATE
 CC OF THE NEURAL TUBE.
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT STAGE 17 DURING THE
 CC INITIATION OF LIMB BUD FORMATION. FROM THAT POINT ONWARDS, THE
 CC EXPRESSION PATTERN EXACTLY MATCHES THE LOCATION OF THE ZONE OF
 CC POLARIZING ACTIVITY (ZPA).
 CC -1- INDUCTION: BY RETINOID ACID.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L28099; AAA72428.1; -.
 CC HSSP: O62226; IVNH.
 CC MEROPS: C46.001; -.
 CC InterPro: IPR000320; HH_signal.
 CC InterPro: IPR001767; Hedgehog_hint.
 CC InterPro: IPR003586; Hedgehog_hintC.
 CC InterPro: IPR003587; Hedgehog_hintN.
 CC InterPro: IPR002203; Intein.
 CC InterPro: IPR001657; SonichH.
 CC Pfam: PF01079; Hint; 1.
 CC Pfam: PF01085; HH_signal; 1.
 CC PRINTS: PR00632; SONICHOG.
 CC PRODOM: PD003042; HH_signal; 1.
 CC SMART: SM00305; HintC; 1.
 CC SMART: SM00306; HintN; 1.
 CC PROSITE: PSS0817; INTEIN_N_TER; 1.
 CC Developmental Protein: Autocatalytic cleavage; Hydrolase; Protease;
 CC Signal: Lipoprotein; Palmitate.
 CC SIGNAL 1 26
 CC CHAIN 27 425
 CC CHAIN 27 200
 CC CHAIN 27 425
 CC SITE 200 201
 CC SITE 246 246
 CC SITE 270 270
 CC ACT_SITE 273 273
 CC BINDING 200 200

POTENTIAL.
 SONIC HEDGEHOG PROTEIN.
 SONIC HEDGEHOG PROTEIN N-PRODUCT.
 SONIC HEDGEHOG PROTEIN C-PRODUCT.
 CLEAVAGE (AUTO-) (PROBABLE).
 INVOLVED IN CHOLESTEROL TRANSFER (BY
 SIMILARITY).
 INVOLVED IN AUTO-CLEAVAGE (BY
 SIMILARITY).
 ESSENTIAL FOR AUTO-CLEAVAGE (BY
 SIMILARITY).
 CHOLESTEROL (BY SIMILARITY).

FT DOMAIN 390 393 POLY-THR.
CC LIPID 27 PALMITATE (BY SIMILARITY).
CC SEQUENCE 425 AA; 46474 MW; DA9627443D4A0173 CRC64;

Query Match 78.1%; Score 1788; DB 1; Length 425;
Best Local Similarity 80.2%; Pred. No. 3e-136;
Matches 353; Conservative 22; Mismatches 43; Indels 22; Gaps 6;

QY 2 LLLARCFVLVIASSLLVCPGLACGPGRCGRRPKKLTPLAYKQFIPNVAEKLGLASG 61
DB 4 MLLRLVLLVGFICALVSSGLTCGPRGICGRHRPKKLTPLAYKQFIPNVAEKLGLASG 63
QY 62 REGGKTRNSEFKELTPYNPDIIFKDEBNGADRLMTQKCKDLMALAISSVMQMPGV 121
DB 64 REGGKTRNSEFKELTPYNPDIIFKDEBNGADRLMTQKCKDLMALAISSVMQMPGV 123
QY 122 RLRTVEGMDHSESLHYEGRAVDITTSDDRSKYGMLARLVEAGFDVWYESSKAH 181
DB 124 KLRTVEGMDHSESLHYEGRAVDITTSDDRSKYGMLARLVEAGFDVWYESSKAH 183
QY 182 IHCYKAENSVAAGSGCGPGSATVHLDEGGTKLVKDLRPGDRLAADOGRLVSDFLT 241
DB 184 IHCYKAENSVAAGSGCGPGSATVHLDEGGTKLVKDLRPGDRLAADOGRLVSDFLT 243
QY 242 FLDRBEGAKKVVYIETLEPRERLLITAAHLFVAP-HNDS--GPTPGPSLFRSRAP 297
DB 244 FLDRBEGAKKVVYIETLEPRERLLITAAHLFVAP-HNDS--GPTPGPSLFRSRAP 302
QY 298 GORVYVVAERGGDRLLPAVAHVTLREBEGAVAPLTAHGTLINRLVASCYVIEEHS 357
DB 303 GORVYVVAERGGDRLLPAVAHVTLREBEGAVAPLTAHGTLINRLVASCYVIEEHS 360
QY 358 WAHRAFAFRLAHLAALAPARTDGGGSGIPAAQSATGARGAFTAGIHWYSQLYHI 417
DB 361 WAHRAFAFRLAHLAALAPARTDGGGSGIPAAQSATGARGAFTAGIHWYSQLYHI 405
QY 418 GFWLDSERMLHPLGMAVKSS 437
DB 406 GFWLDSERMLHPLGMAVKSS 425

RESULT 5
SHH_CYNPY STANDARD: PRT: 432 AA.
AC 090385;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sonic hedgehog protein precursor (SHH).
GN SHH.
OS Cynops pyrrhogaster (Japanese common newt).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroides; Salamandridae; Cynops.
OX NCBI_TaxID=8330;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=96136334; PubMed=8573168;
RA Takabatake T., Takahashi T.C., Inoue K., Ogawa M., Takeshima K.;
RT "Activation of two Cynops genes, fork head and sonic hedgehog, in animal cap explants.";
RL Blochem. Biophys. Res. Commun. 218:395-401(1996).
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED IN LIMB FORMATION, PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND VENTRAL SOMITE DIFFERENTIATION. INDUCES ECTOPIC CEMENT GLAND FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESENTS THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE

CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- INDUCTION: ACTIVATED BY ACTININ, BASIC FIBROBLAST GROWTH FACTOR (BFGF) AND FORK HEAD.
CC -1- PIM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: D63339; BAA09657.1; -
DR HSSP: Q62226; 1VH.
DR MEROPS: C46.002; -
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR00167; Hedgehog_hint.
DR InterPro: IPR003587; Hedgehog_hintc.
DR InterPro: IPR002203; Intein.
DR InterPro: IPR001657; SonicHH.
DR Pfam: PF01079; Hint; 1.
DR PRINTS: PR00632; SONICHHOG.
DR PRODOM: PD003042; HH_signal; 1.
DR SMART: SM00305; Hintc; 1.
DR SMART: SM00306; Hintc; 1.
DR PROSITE: PS50817; INTEIN_N_TER; 1.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW signal; Lipoprotein; Palmitate.
FT SIGNAL 1 26
FT CHAIN 27 432
FT CHAIN 27 200
FT CHAIN 201 432
FT SITE 200 201
FT SITE 268 268
FT ACT_SITE 271 271
FT BINDING 200 200
FT LIPID 27 27
FT SEQUENCE 432 AA; 47847 MW; BA55C7E746C8B5A8 CRC64;

Query Match 67.4%; Score 1542.5; DB 1; Length 432;
Best Local Similarity 70.2%; Pred. No. 1.7e-116;
Matches 313; Conservative 32; Mismatches 74; Indels 27; Gaps 7;

QY 2 LLLARCFVLVIASSLLVCPGLACGPGRCGRRPKKLTPLAYKQFIPNVAEKLGLASG 61
DB 4 MLLRLVLLVGFICALVSSGLTCGPRGICGRHRPKKLTPLAYKQFIPNVAEKLGLASG 63
QY 62 REGGKTRNSEFKELTPYNPDIIFKDEBNGADRLMTQKCKDLMALAISSVMQMPGV 121
DB 64 REGGKTRNSEFKELTPYNPDIIFKDEBNGADRLMTQKCKDLMALAISSVMQMPGV 123
QY 122 RLRTVEGMDHSESLHYEGRAVDITTSDDRSKYGMLARLVEAGFDVWYESSKAH 181
DB 124 KLRTVEGMDHSESLHYEGRAVDITTSDDRSKYGMLARLVEAGFDVWYESSKAH 183
QY 182 IHCYKAENSVAAGSGCGPGSATVHLDEGGTKLVKDLRPGDRLAADOGRLVSDFLT 241
DB 184 IHCYKAENSVAAGSGCGPGSATVHLDEGGTKLVKDLRPGDRLAADOGRLVSDFLT 243

Oy	242	FLDDECAKKVEYIEFLPEPRRELLTAALLVPA----	PHN0SG-----	PTP0SALF	291
Db	244	FMDKEETVRAKVEYIEET--SREVRLLTAALLLEVG	AIHPNDGGGDFRSVFSAGRSKF		301
Oy	292	ASRRPQGVRYVAENRGCDRRLLPAAVHSVTLRE	EAGAVAPLTAHGTLLIRNVLASCYA	351	
Db	302	ASSVRAGHRYLVTVDRGEG--RGLEAEATVEERYLL	EEANTGAAPATVTAAGTVIDRVLASCYA	358	
Oy	352	VIEESHVNAHRAFPFRLAHLAALLAAPARTDGGGGS	IPAA0SATEARCAEPTACIHWYS	411	
Db	359	VIEESHVNAHMAFAFLRAGLGLISFFSPD0YSHS--	PPASQ0SE-----GVHWYS	406	
Oy	412	QLLYHGMWLDSETHPLCLMAVKSX	437		
Db	407	EILRIGTWIYQEDTHPLCLMAKXS	432		
RESULT 6					
SHH_XENLA					
ID	SHH_XENLA	STANDARD;	PRT;	444	AA.
AC	092000;	091894;			
DT	15-JUL-1999	(Rel. 38, Created)			
DT	15-OCT-1999	(Rel. 38, Last sequence update)			
DT	16-JUL-2001	(Rel. 40, Last annotation update)			
DE	SHH.	"Xenopus hedgehog protein precursor (X-SHH) (VH-1).			
GN	SHH.				
OS	Xenopus laevis (African clawed frog).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;				
OC	Xenopodidae; Xenopus.				
OX	NCBI_TaxID=8355;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Intestine;				
RX	MEDLINE=95357169;	PubMed=7630736;			
RA	Stolow M.A., Shi Y.-B.;				
RT	"Xenopus sonic hedgehog as a potential morphogen during embryogenesis				
RT	and thyroid hormone-dependent metamorphosis.";				
RL	Nucleic Acids Res. 23:2555-2562(1995).				
RL	[2]				
RN	SEQUENCE FROM N.A.				
RP	TISSUE=Embryo;				
RC	MEDLINE=95401852;	PubMed=7671800;			
RX	Ecker S.C., McGrew L.L., Lal C.-J., Lee J.J., von Kessler D.P.,				
RA	Moon R.T., Beachy P.A.;				
RT	"Distinct expression and shared activities of members of the hedgehog				
RT	gene family of Xenopus laevis.";				
RL	Development 121:2337-2347(1995).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Notochord;				
RX	MEDLINE=96028338;	PubMed=7551564;			
RA	Ruiz i Altaba A., Jessell T.M., Roelink H.;				
RT	"Restrictions to floor plate induction by hedgehog and winged-helix				
RT	genes in the neural tube of frog embryos.";				
RL	Mol. Cell. Neurosci. 6:106-121(1995).				
CC	-1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF				
CC	PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED				
CC	IN LIMB FORMATION, PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND				
CC	VENTRAL SOMITE DIFFERENTIATION. INDUCES ECOTOPIC CEMENT GLAND				
CC	FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH				
CC	FUNCTIONS IN ASSOCIATION WITH SMOOTHEDEN (SMO). TO ACTIVATE THE				
CC	TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC				
CC	REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY				
CC	SMIIMILARITY).				
CC	-1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE				
CC	CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE				
CC	CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM				
CC	TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).				
CC	-1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN NOTOCHORD AND NEURAL				
CC	FLOOR PLATE DURING EMBRYOGENESIS. IN TADPOLE, HIGH EXPRESSION IS				
CC	OBSERVED IN PANCREAS/STOMACH, MODERATE EXPRESSION IN TAIL, AND LOW				

	CC	EXPRESSIN IN INTESTINE, BRAIN, AND HIND LIMB.
	CC	-1- DEVELOPMENTAL STAGE: FIRST DETECTED AT THE NEURULA (STAGES 16-17).
	CC	FIRST PEAK OF EXPRESSION AROUND TADPOLE HATCHING (STAGES 33-40).
	CC	HIGH EXPRESSION OBSERVED IN INTESTINE AT THE CLIMAX OF MORPHOGENESIS (STAGES 60-62) WHEN INTESTINE EPITHELIAL UNDERGOES MORPHOGENESIS.
	CC	-1- INDUCTION: BY THYROID HORMONE.
	CC	-1- PPM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
	CC	-1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
	CC	-----
	CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
	CC	-----
	DR	EMBL; L39213; AAC42227.1; .-
	DR	EMBL; U26314; AAA85162.1; .-
	DR	EMBL; L35248; AAA49981.1; .-
	DR	HSSP; O62226; 1VHH.
	DR	MEROPS; C46.002; .-
	DR	InterPro; IPR000320; HH_signal.
	DR	InterPro; IPR001767; Hedgehog_hntc.
	DR	InterPro; IPR003586; Hedgehog_hntc.
	DR	InterPro; IPR003587; Hedgehog_hntn.
	DR	InterPro; IPR002203; Intein.
	DR	InterPro; IPR001657; SonicHh.
	DR	Pfam; PF01079; Hntc; 1.
	DR	Pfam; PF01085; HH_signal; 1.
	DR	PRINTS; PR00632; SONICHOG.
	DR	ProDom; PD003042; HH_signal; 1.
	DR	SMART; SM00305; Hntc; 1.
	DR	SMART; SM00306; Hntn; 1.
	DR	PROSITE; PS50817; INTEIN_N_TER; 1.
	RW	Developmental protein; Autocatalytic cleavage; Hydrolyase; Protease; Signal; Lipoprotein; Palmitate; Repeat.
	KM	SIGNAL
	FT	1 24 POTENTIAL.
	FT	CHAIN 25 444 SONIC HEDGEHOG PROTEIN.
	FT	CHAIN 25 198 SONIC HEDGEHOG PROTEIN N-PRODUCT.
	FT	CHAIN 199 444 SONIC HEDGEHOG PROTEIN C-PRODUCT.
	FT	SITE 199 199 CLEAVAGE (AUTO-) (BY SIMILARITY).
	FT	SITE 266 266 INVOLVED IN AUTO-CLEAVAGE (BY SIMILARITY).
	FT	ACT_SITE 269 269 ESSENTIAL FOR AUTO-CLEAVAGE (BY SIMILARITY).
	FT	DOMAIN 386 409 3 X 8 AA TANDEM REPEATS OF Q-V-D-L-Q-S-H-H.
	FT	REPEAT 386 393 1.
	FT	REPEAT 394 401 1.
	FT	REPEAT 403 409 3.
	FT	BINDING 198 198 CHOLESTEROL (BY SIMILARITY).
	FT	LIPID 25 25 PALMITATE (BY SIMILARITY).
	FT	CONFLICT 5 9 TOSIL -> NSNLCK (IN REF. 3).
	FT	CONFLICT 302 319 DPKMTLVANVEKVDLE -> ESODHDLBSGRKKRRILLR (IN REF. 3).
	FT	N -> S (IN REF. 3).
	SO	SEQUENCE 444 AA; 49453 MW; 73B4E933FA2EEFF2 CRC64;
		Query Match 66.6%; Score 1525.5; DB 1; Length 444;
		Match Local Similarity 67.4%; Pred. No. 4e-115;
		Batches 304; Conservative 38; Mismatches 88; Indels 21; Gaps 6;
1	MILLARCFVLIVLSSLLVCGLACGPGRCGRKRNNPKRLTPLAYKQFIPIPVAAETKLGA5 60	

```

DB      1  MVAATOSLLSLFICLTPPCACGPGRGICRRRPRKLTPLATYKOFIPNVAEKTIGAS 60
OY      61  GREGKITNSERKELETNYNDPIEFKDEMTGADRLTORCKDKLNLATISVMNQMG 120
DB      61  GREGKITNSOCFKELTNYNDPIEFKDEESTGADRLTORCKDKLNLATISVMNQMG 120
OY      121  VALRTVEGDEDOGHSEESLHYEGRAVDITTSDRSKYGMRLARLAVEGFMVVEESKA 180
DB      121  VKLRVTEGDEDOGHSEESLHYEGRAVDITTSDRSKYGMRLARLAVEGFMVVEESKA 180
OY      181  HHCSYKAANSVAASGCGFPGSATYHLEOGCTKLYKDLRPGDRVLAADOGRLYSDFL 240
DB      181  HHCSYKAANSVAASGCGFPGSATYHLEOGCTKLYKDLRPGDRVLAADOGRLYSDFL 240
OY      241  TELDEGACAKFEVYIETLEPERERLLTAHLLEFVAPHNDSGPTSPALSRVPRGOR 300
DB      241  MLDGDRDVKKLFYIET--SQRKITLTAHLLEFVAGTQVNG--TRSKYVFSNIQPGDL 297
OY      301  VYVVAERGGDRLLPAAVHSVTLREBEAGAVAPLTAHGTILINRLVASCYAVIEHSHMAH 360
DB      298  IYTADEK--TMTLKAVKVEKVDL-EEDTGAVAPLTAHGTIVIDQVLASCYAVIEEHTMAH 354
OY      361  RAPAPRLAHALLAARARTDGGCGGSTRPAQ-----SATFARGAPPTAG 406
DB      355  LAFAPLRFEGMSLSSYIYR-RDSSPSPGLOPHHQVDLQSHHOVDLQSHHQVDLQSHHQVDL 413
OY      407  IHMYSQLYHIGTWLIDSETMHPGLMAVYKSS 437
DB      414  IHMYSQLYHIGTWLIDSETMHPGLMAVYKSS 444

```

```

RESULT 7
SHH_BRARE
AC      092008: 013170: 013171:
ID      15-JUL-1998 (Rel. 38, Created)
DT      15-JUL-1999 (Rel. 38, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Sonic hedgehog protein precursor (SHH) (VNH-1).
GN      SHH OR VNH1.
OS      Brachydanio rerio (zebrafish) (Danio rerio).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC      Cyprinidae; Danio.
OX      NCBI_TaxID=7955;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=embryo;
RX      MEDLINE=94170375; PubMed=8124714;
RA      Roelink H., Augsburg A., Heemskerk J., Korzh V., Nollin S.,
RA      Ruiz I Altaba A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,
RA      Dodd J.;
RT      "Floor plate and motor neuron induction by vhh-1, a vertebrate homolog
RT      of hedgehog expressed by the notochord.";
RL      Cell 76:761-775(1994).
RN      [2]
RP      SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
RX      MEDLINE=96014264; PubMed=7583153;
RA      Ekker S.C., Ungar A.R., Greenstein P., von Kessler D.P., Porter J.A.,
RA      Moon R.T., Beachy P.A.;
RT      "Patterning activities of vertebrate hedgehog proteins in the
RT      developing eye and brain.";
RL      Curr. Biol. 5:944-955(1995).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=96083328; PubMed=7579523;
RA      Fietz W.J., Concordet J.-P., Barbosa R., Johnson R., Krauss S.,
RA      McMahon A.P., Tabin C., Ingham P.W.;
RT      "The hedgehog gene family in Drosophila and vertebrate development.";
RL      Development Suppl. 43-51(1994).
RN      [4]
RP      SEQUENCE FROM N.A.

```

```

RX      MEDLINE=99225458; PubMed=10207136;
RA      Muller F., Chang B., Albert S., Fischer N., Tora L., Strahle U.;
RT      "Intronic enhancers control expression of zebrafish sonic hedgehog in
RT      floor plate and notochord.";
RL      Development 126:2103-2116(1999).
RN      [5]
RP      SEQUENCE OF 30-92 AND 113-170 FROM N.A.
RC      TISSUE=Muscle;
RX      MEDLINE=97075114; PubMed=8917540;
RA      Zardoya R., Abouheif E., Meyer A.;
RT      "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
RT      closely related to the zebrafish.";
RL      Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC      -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
CC      PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
CC      NOTOCHORD THAT INDUCES SOMITE PATTERNING. DORSO-VENTRAL PATTERNING
CC      OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS
CC      FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC)
CC      RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO
CC      ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,
CC      PTC REPRESENTS THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
CC      SIMILARITY).
CC      -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC      CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC      CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC      TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC      -1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL
CC      TUBE AND BRAIN. ALSO FOUND IN THE NOTOCHORD AND IN DEVELOPING FIN
CC      BUD. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN DOMAINS THAT
CC      INCLUDE A DISCRETE REGION IN THE FLOOR OF THE Diencephalon.
CC      -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE IN THE INNER CELL LAYER OF
CC      THE EMBRYONIC SHIELD DURING GASTRULATION. BY 9.5 HRS OF
CC      DEVELOPMENT, EXPRESSED IN A CONTINUOUS BAND THAT EXTENDS FROM THE
CC      TAIL TO THE HEAD. THE ANTERIOR BOUNDARY OF EXPRESSION BEING
CC      POSITIONED IN THE CENTER OF THE ANIMAL POLE ANTERIOR TO THE
CC      PRESUMPTIVE MIDBRAIN.
CC      -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC      AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC      THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC      A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC      TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC      TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC      OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC      ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC      C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC      -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: L27585; AAA20998.1;
DR      EMBL: U30711; AAC59742.1;
DR      EMBL: Z35669; CA84738.1;
DR      EMBL: AF124382; AAD47913.1;
DR      EMBL: U51351; AAB38575.1;
DR      EMBL: U51370; AAB38593.1;
DR      HSSP: Q62226; IVH.
DR      MEROPS: C46.002;
DR      ZFIN: ZDB-GENE-980526-166; shh.
DR      InterPro: IPR000320; HH_signal.
DR      InterPro: IPR001767; Hedgehog_hint.
DR      InterPro: IPR003586; Hedgehog_hintC.
DR      InterPro: IPR003587; Hedgehog_hintN.
DR      InterPro: IPR002203; Intein.
DR      InterPro: IPR001657; SonicHH.
DR      Pfam: PF01079; Hint; 1.
DR      Pfam: PF01085; HH_signal; 1.
DR      PRINTS: PR00632; SONICHHOG.

```

DR PRODOM: PD003042; HH.signal; 1.
 DR SMART; SM00305; HintC; 1.
 DR SMART; SM00306; HintC; 1.
 DR PROSITE; PSS0817; INTERIN_N_TER; 1.
 KM Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KM Signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 23
 FT CHAIN 24 418
 FT CHAIN 24 197
 FT CHAIN 198 418
 FT SITE 197 198
 FT SITE 243 243
 FT SITE 267 267
 FT ACT_SITE 270 270
 FT BINDING 197 197
 FT LIPID 24 24
 SQ SEQUENCE 418 AA; 46402 MW; CF00DAFFED2F5795 CRC64;
 Query Match 66.3%; Score 1517; DB 1; Length 418;
 Best Local Similarity 67.7%; Pred. No. 1.8e-114;
 Matches 294; Conservative 47; Mismatches 75; Indels 18; Gaps 3;
 QY 4 LLARCPVLVLAASSLLVPGIACGPGRGFRGRRHKKLPLAYKQFIPVNAEKTIGASGRY 63
 DB 3 LLFTVLVSLTLTSLVYSGIACGPGRGFRGRRHKKLPLAYKQFIPVNAEKTIGASGRY 62
 QY 64 EGGTRNSERKELTPYNPDIIFKDEENGADRLMTORCKDKNALALISTMNQPGVRL 123
 DB 63 EGGTRNSERKELTPYNPDIIFKDEENGADRLMTORCKDKNALALISTMNQPGVRL 122
 QY 124 RTEGMEDGHSESLHYEGRAVDITTSDBRSKYGLARLAVEAGFDWYYSKAHII 183
 DB 123 RVTGCMEDGHSESLHYEGRAVDITTSDBRSKYGLARLAVEAGFDWYYSKAHII 182
 QY 184 CSYKAENSVAKSGCGEPGSAIVHLEQGTLYKDLRPGDRLAADDGRLLYSDFLTEL 243
 DB 183 CSYKAENSVAKSGCGEPGSAIVHLEQGTLYKDLRPGDRLAADDGRLLYSDFLTEL 242
 QY 244 DRDGAKKVPIVITLPERRELLTAAHLFVAPRNDGSPGSALEFASVVRGQRYV 303
 DB 243 DRDSTTRRVFVITLTOEVEKILTAAHLFVL-DNSTEDLTHTAAYASVVRGQRYV 301
 QY 304 VAESGGDRRLPAVHVSVTLEEEAGVAPLTAHTILLINVLASCYAVIEHSHARAF 363
 DB 302 VDDSGQLKSVIVQRIYI---EORGSAFAPYTAHTIVVDRIILASCYAVIEDGLAHIAF 357
 QY 364 APFRLAHLAALAPARTDGGGSGIPASASATEARCAEPTAGIHWYSQLLYHIGTWLID 423
 DB 358 APALRYVYVSFLRP-----QNSSRSNATLQEGGVHWYSRLLYQGTWILD 404
 QY 424 SETMHPLGMAVKSS 437
 DB 405 SNMLPLGMSVNSS 418
 RESULT 8
 TMH_BRARE STANDARD; PRT; 416 AA.
 ID TMH_BRARE
 AC Q90419;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Tiggy-winkle hedgehog protein precursor (TMH).
 GN Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NC NCBI_TaxID=7955;
 RN [1]

RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
 RC TISSUE-Embryo;
 RX MEDLINE-96014264; PubMed-7581133;
 RA Ekker S.C., Ungar A.R., Greenstein P., von Kessler D.P.,
 RA Porter J.A., Moon R.T., Beachy P.A.,
 RT "Patterning activities of vertebrate hedgehog proteins in the
 RT developing eye and brain."
 RL Curr. Biol. 5:944-955(1995).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. INVOLVED IN DORSO-VENTRAL
 CC PATTERNING OF THE BRAIN AND IN EARLY PATTERNING OF THE DEVELOPING
 CC EYES.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL
 CC TUBE AND BRAIN. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN
 CC DOMAINS THAT INCLUDE A DISCRETE REGION IN THE FLOOR OF THE
 CC DIENCEPHALON. NOT DETECTED IN THE NOTOCHORD OR DEVELOPING FIN BUD.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U30710; AAC59741.1; -.
 CC HSSP: Q62226; 1VHH.
 DR MEROPS; C46.001; -.
 DR ZFIN; ZDB-GENE-980526-41; tmh.
 DR InterPro; IPR000320; HH_signal.
 DR InterPro; IPR001767; Hedgehog_hntc.
 DR InterPro; IPR003586; Hedgehog_hntc.
 DR InterPro; IPR003587; Hedgehog_hntn.
 DR InterPro; IPR001657; Sonichn.
 DR Pfam; PF01079; HintC; 1.
 DR Pfam; PF01085; HH_signal; 1.
 DR PRINTS; PR00632; SONICHOG.
 DR PRODOM; PD003042; HH_signal; 1.
 DR SMART; SM00305; HintC; 1.
 DR SMART; SM00306; HintC; 1.
 KM Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KM Signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 26
 FT CHAIN 27 416
 FT CHAIN 201 416
 FT SITE 200 201
 FT SITE 270 270
 FT ACT_SITE 273 273
 FT BINDING 200 200
 FT LIPID 27 27
 SQ SEQUENCE 416 AA; 46576 MW; 61EC2218309CFE59 CRC64;
 Query Match 64.0%; Score 1464; DB 1; Length 416;
 Best Local Similarity 66.7%; Pred. No. 3.2e-110;
 Matches 287; Conservative 43; Mismatches 64; Indels 36; Gaps 6;

QY 359 AHRFAFRLAHLAALAPRTDGGGGSIPAAQSATEANGAPPTAGIHWYSQLYHIC 418
 DB 354 AQMAFWRLYHSL-----GG-PGVGG-----DGVHWYSGLYRLG 389
 QY 419 TWLDETHMPLG 431
 DB 390 RMLLPDSFRLG 402
 RESULT 10
 ID IHH_HUMAN STANDARD; PRT; 411 AA.
 AC Q14623; O43322;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Indian hedgehog protein precursor (IHH) (HHG-2).
 GN IHH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tate G., Kishimoto K., Mitsuura T.;
 RT "Expression of Sonic hedgehog and its receptor Patched/Smoothed in
 human cancer cell lines and embryonic organs.";
 RL J. Biochem. Mol. Biol. Biophys. 4:27-34(2000).
 RN [2]
 RP SEQUENCE OF 100-411 FROM N.A.
 RC TISSUE=Fetal lung;
 RX MEDLINE=96070431; PubMed=7590746;
 RA Marigo V., Roberts D.J., Lee S.M.K., Teukurov O., Levi T.,
 RA Gattler J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seldman C.E.,
 RA Jankins N.A., Seldman J.G., McMahon A.P., Tabin C.;
 RT "Cloning, expression, and chromosomal location of SHH and IHH: two
 human homologues of the Drosophila segment polarity gene hedgehog.";
 RL Genomics 28:44-51(1995).
 RN [3]
 RP SEQUENCE OF 124-172 FROM N.A.
 RX MEDLINE=95236997; PubMed=7720571;
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandi B.K.,
 RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;
 RT "Products, genetic linkage and limb patterning activity of a murine
 hedgehog gene.";
 RL Development 120:3339-3353(1994).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)
 CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO
 CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN
 CC ENDOCHONDRAL OSSIFICATION. MAY REGULATE THE BALANCE BETWEEN GROWTH
 CC AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION
 CC OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP) (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN EMBRYONIC LUNG, AND IN ADULT
 CC KIDNEY AND LIVER.
 CC -1- PPM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB018076; BAA33523.2; -
 DR EMBL; AB010092; BAA33523.2; JOINED.
 DR EMBL; AB018075; BAA33523.2; JOINED.
 DR EMBL; L38517; AAA62178.1; -
 DR HSSP; 062226; 1VHH.
 DR MEROPS; C46.003; -
 DR Genew; HGNC:5956; IHH.
 DR MIM; 600726; -
 DR InterPro; IPR000320; HH_signal.
 DR InterPro; IPR001767; Hedgehog_hnt.
 DR InterPro; IPR003586; Hedgehog_hntC.
 DR InterPro; IPR003587; Hedgehog_hntN.
 DR InterPro; IPR002203; Intein.
 DR Pfam; PF01079; Hnt; 1.
 DR Pfam; PF01085; HH_signal; 1.
 DR SMART; SM00305; HntC; 1.
 DR SMART; SM00306; HntN; 1.
 DR PROSITE; PSS0817; INTEIN_N_TER; 1.
 DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 27
 FT CHAIN 28 411
 FT CHAIN 28 202
 FT CHAIN 203 411
 FT SITE 202 203
 FT SITE 248 248
 FT SITE 272 272
 FT ACT_SITE 275 275
 FT BINDING 202 202
 FT LIPID 28 28
 FT CONFLICT 100 100
 FT CONFLICT 246 246
 FT CONFLICT 309 309
 FT SEQUENCE 411 AA; 4518 MW; 14C904A3C344DBA CRC64;
 SQ
 Query Match 55.5%; Score 1271.5; DB 1; Length 411;
 Best Local Similarity 60.6%; Pred. No. 9.4e-95;
 Matches 260; Conservative 40; Mismatches 92; Indels 37; Gaps 8;
 QY 8 CELVTLASSLIVCPGL-ACGPGRGFG-KRRHPKLTPLAYKQFTPNVAEKTLAGSGRYEG 65
 DB 13 CLVTL--LLVVPAAAGCGGVRVGSRRRPRLVPLAYKQFSPNVEKTLAGSGRYEG 69
 QY 66 KITNSRREKLTTPNYNDITFEKDEMTGADRLTORCKCKLNLAISVNNQMGVRLV 125
 DB 70 KITNSRREKLTTPNYNDITFEKDEMTGADRLTORCKCKLNLAISVNNQMGVRLV 129
 QY 126 TEGWDEGHSEESLHAEGRVADITTSDRSKXGKMLARLAVEGFPVYVESRAHICS 185
 DB 130 TEGWDEGHSEESLHAEGRVADITTSDRSKXGKMLARLAVEGFPVYVESRAHICS 189
 QY 126 TEGWDEGHSEESLHAEGRVADITTSDRSKXGKMLARLAVEGFPVYVESRAHICS 185
 DB 130 TEGWDEGHSEESLHAEGRVADITTSDRSKXGKMLARLAVEGFPVYVESRAHICS 189
 QY 186 VKAENVAASGCGPESATVHLEGGTKLVKDLRPGDRLVLAADQGRLLYSDFLEFLDR 245
 DB 190 VKSHSAAAKGCGFPAGACVRLSGARVALSAVRPDRVLAEMEDSPFSFVLLIDR 249
 QY 246 DEGAKKFEVETLEPERRLITTAHLLFVAPHDSPTGFSALFASRVPRGGRVYVA 305
 DB 250 EPHRLRAFVYETODPPRLALTPAHLLFTA-DHTEPARAFRTFASHVQGYVLVAG 308
 QY 306 ERGDRRLPAVAHVSVTLREEEAGAVAPLTANGTILINRVLASCAVYIEESHNAHRAFP 365
 DB 309 APG---LQPARVAAS-TYHALGAVAPLTRKHGLTVEDVVASCFAAVDHHLAQLAFWP 363
 QY 366 FRLAHLAALAPARTDGGGGSIPAAQSATEARGA-EPPTAGIHWYSQLYHICGTLWLD 424

Db 252 NLRAPVETEDPPRLATPLAHLFLA-DNTERPAHFRATASHVQGVYVSGVP 310
 Qy 308 GGDRLPLPAVHVTLEEEAGAYAPLTANGTILINRYLASCVAVIEHSHMAHRAFAFR 367
 Db 311 LQPARVAAYS-THVALGSAFLRGTLLVEDVASCFAVAADHHLAQLAFNPLR 365
 Qy 368 LAHALLAALAPARTDGGGGSIPAQSAATEARGA-EPPAGIHWYSQLLYHIGTWLDSET 426
 Db 366 LFPSTL-----AMGSMTPEGVHWPQMLYRLGRLLLEST 400
 Qy 427 MHPILGM 433
 Db 401 FHPILGMS 407
 RESULT 12
 IHH_XENLA STANDARD; PRT; 409 AA.
 AC Q91612;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Indian hedgehog protein precursor (IHH) (banded hedgehog protein) (X-BHH).
 GN IHH OR BHH.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=95401852; Pubmed=7671800;
 RA Ekker S.C., McGrew L.L., Lai C.-J., Lee J.J., von Kessler D.P.,
 RA Moon R.T., Beachy P.A.;
 RT "Distinct expression and shared activities of members of the hedgehog gene family of *Xenopus laevis*."
 RL Development 121:2337-2347(1995).
 CC -1- FUNCTION: SIGNAL INVOLVED IN THE EARLY INDUCTION AND PATTERNING OF ANTERODORSAL ECTODERM, NERVOUS SYSTEM AND SOMITES. INDUCES ECTOPIC CEMENT GLAND FORMATION IN EMBRYOS.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. HEDGEHOG PROTEIN IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- DEVELOPMENTAL STAGE: DETECTABLE WITHIN THE EARLY GASTRULA. AT STAGE 14 NEURULA, HIGH EXPRESSION IN NEUROECTODERM. EXPRESSED THROUGHOUT THE NEURAL PLATE AND SUBSEQUENTLY IN BOTH THE NERVOUS SYSTEM AND IN THE DERMATOME OF SOMITES.
 CC -1- INDUCTION: BY ACTIVIN.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: U26404; AAA85165.1; -.
 DR HSSP; Q62226; LVHH.

DR MEROPS; C46.0PW; -.
 DR InterPro: IPR000320; HH_signal.
 DR InterPro: IPR001767; Hedgehog_hintc.
 DR InterPro: IPR003586; Hedgehog_hintc.
 DR InterPro: IPR003587; Hedgehog_hintc.
 DR InterPro: IPR001657; SonICHH.
 DR Pfam: PF01079; Hh_signal.1.
 DR Pfam: PF01085; Hh_signal.1.
 DR PRINTS: PR00632; SONICHOOG.
 DR ProDom: PD003042; HH_signal.1.
 DR SMART: SM00305; Hintc.1.
 DR SMART: SM00306; Hintc.1.
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 23
 FT CHAIN 24 409
 FT CHAIN 24 197
 FT CHAIN 198 409
 FT SITE 197 198
 FT SITE 267 267
 FT ACT_SITE 270 270
 FT BINDING 197 197
 FT DOMAIN 32 35
 FT LIPID 24 24
 FT SEQUENCE 409 AA; 45591 MW; 6B265367FB98627 CRC64;

Query Match 55.48; Score 1267; DB 1; Length 409;
 Best Local Similarity 59.94; Pred. No. 2, 1e-94;
 Matches 257; Conservative 49; Mismatches 95; Indels 28; Gaps 6;

Qy 5 LARCEVLIALSLVPCGL-ACGPGRGFRRRHKTLPLAYQGFIVNAEKTGASGRY 63
 Db 3 LPRVLLCAHALLSGAVKCGGRVYGRRRRTKLSPLSTYKQSPNVEPKTGASGRY 62
 Qy 64 EGRITRNSERKELTPYNDIIFKDEENTGADRLMTORCKDKLNALASVMNQMPVRL 123
 Db 63 EGRISRSERKELTPYNDIIFKDEITGADRLMTORCKDKLNALASVMNQMPVKL 122
 Qy 124 RYEGWDEGHSESLHYEGRAVDITTSRDSKCYMLARLAVAGFDWYVESKAHII 183
 Db 123 RYEGWDEGHSESLHYEGRAVDITTSRDRKKCYMLARLAVAGFDWYVESKAHII 182
 Qy 184 CSVAEKSVAKSGCGPGSATVHLEGGKTYDLPGRVYLAADOGRLVYSDFTFL 243
 Db 183 CSVAEKSAAKTGGCGPGGALATLSEGEKIPVSQLPLAMDNSGRPTYSDFLFL 242
 Qy 244 DRDEGAKKVVYETLEPRERLLTAHLFLVAPANDSGPTPGPSALFASVRPGQHYVY 303
 Db 243 DHSPEKHEMQVITQDPRHRLFLTPAHLFLVS-DNSTPASEQAVFASVRRGQYILV 301
 Qy 304 VAEGGRRLLPAVHVTLEEEAGAYAPLTANGTILINRYLASCVAVIEHSHMAHRA 363
 Db 302 SNVVG----LIPAKVRVSN-TQTYGAYAPLTQGTIVDVVASCALVQOKOQADIVY 356
 Qy 364 APFLAHLAALAPARTDGGGGSIPAQSAATEARGA-EPPAGIHWYSQLLYHIGTWLD 423
 Db 357 WPLRVLYNL-----GIAGTQPSQ-----MGHWYSKALYHIGRLIHL 395
 Qy 424 SETMHPILGM 432
 Db 396 GNEFHPILGI 404
 RESULT 13
 IHH_BRARE STANDARD; PRT; 412 AA.
 AC Q98862;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Indian hedgehog protein precursor (IHH) (Fchidna hedgehog protein).

GN IHH OR EHH.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NCBI_TaxID=7955;
 RN
 RP SEQUENCE FROM N.A.
 RA MEDLINE=96310864; PubMed=8684485;
 RT Currie P.D., Ingham P.W.;
 "Induction of a specific muscle cell type by a hedgehog-like protein
 in zebrafish".
 RL Nature 382:452-455(1996).
 RN
 RP SEQUENCE OF 113-170 FROM N.A.
 RC TISSUE=Muscle;
 RX MEDLINE=97075114; PubMed=8917540;
 RA Zardoya R., Abouheif E., Meyer A.;
 "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
 closely related to the zebrafish".
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
 NOTOCHORD THAT INDUCES SOMITE PATTERNING AND MUSCLE PIONEER
 DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN THE NOTOCHORD.
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT THE MID-GASTRULA STAGE.
 CC DISAPPEARS AT THE END OF THE SOMITOGENESIS.
 CC -1- P.M.: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 or send an email to license@sdb-sdb.ch).
 CC
 DR EMBL: Y08426; CAA69702.1; .
 DR HSSP: 062226; 1VHH.
 DR MEROPS: C46 0PW.
 DR ZFIN: ZDB-GENE-980526-135; eh.
 DR InterPro: IPR000320; HH_signal.
 DR InterPro: IPR001767; Hedgehog_hinc.
 DR InterPro: IPR003586; Hedgehog_hinc.
 DR InterPro: IPR003587; Hedgehog_hinc.
 DR InterPro: IPR002203; Inteln.
 DR InterPro: IPR001657; Sonichn.
 DR Pfam: PF01079; Hinc_1.
 DR Pfam: PF01085; HH_signal_1.
 DR PRINTS: PR00632; SONICHOG.
 DR PRODOM: PD003042; HH_signal_1.
 DR SMART: SM00305; Hinc_1.
 DR SMART: SM00306; Hinc_1.
 DR PROSITE: PS50817; INTEIN_N_TER_1.
 KM Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 RN Signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 412 INDIAN HEDGEHOG PROTEIN.
 FT 197 INDIAN HEDGEHOG PROTEIN N-PRODUCT.

FT CHAIN 198 412 INDIAN HEDGEHOG PROTEIN C-PRODUCT.
 FT SITE 197 198 CLEAVAGE (AUTO-) (BY SIMILARITY).
 FT SITE 246 246 INVOLVED IN CHOLESTEROL TRANSFER (BY
 FT SITE 269 269 INVOLVED IN AUTO-CLEAVAGE (BY
 FT ACT_SITE 272 272 SIMILARITY).
 FT BINDING 197 197 ESSENTIAL FOR AUTO-CLEAVAGE (BY
 FT LIPID 24 24 CHOLESTEROL (BY SIMILARITY).
 FT CONFLICT 121 121 PALMITATE (BY SIMILARITY).
 FT SEQUENCE 412 AA: 45746 MW: 21043 F052A278CAL CRC64:
 Query Match 53.6%; Score 1228; DB 1; Length 412;
 Best Local Similarity 60.4%; Pred. No. 2,9e-91;
 Matches 252; Conservative 43; Mismatches 80; Indels 42; Gaps 8;
 QY 25 CGPGRGFGKRRHPKKTPLAYKOFITNNAEKTIGASRGKTRNSERFKELTPYND 84
 DB 24 CGPGRGFGKRRHPKKTPLAYKOFITNNAEKTIGASRGKTRNSERFKELTPYND 83
 QY 85 IIFKDEENTGADRLMTORCKDKLNALISVMNOMPVRRLVTEGMDGDHSESLHYEG 144
 DB 84 IIFKDEENTGADRLMTORCKDKLNALISVMNOMPVRRLVTEGMDGDHSESLHYEG 143
 QY 145 RAVDITTSRDRSKYKMLARLAVEAGFDWYVESKAHICSVKAENSVAAKSGCGPESA 204
 DB 144 RAVDITTSRDRSKYKMLARLAVEAGFDWYVESKAHICSVKAENSVAAKSGCGPESA 203
 QY 205 TVHLEGGTRKLVKDRPGRVLA---ADDOGRLLSYDFLTLDROGAKKRVYETLTP 261
 DB 204 LVYMKDSHROLDLQAGLVLAESGSDGTGLDYSEVLTFDRRPTQKHVYIRT-ED 262
 QY 262 REBLITTAHLFLVAPHNDSG---PDPGP-SALFASRVPRGQVYVVAARGDRLPLPA 317
 DB 263 GASVSLTAHLFLKRVGNKSNKEPFGAVRTIFASDAVQGLL-----GKLRFRSQ 317
 QY 318 VHSVTLREEGAYAPLTAHTGILLIRVLASCAVAYEESHMAHRAFPRLAHALLALA 377
 DB 318 ITHVGRREQ-GIYPLPLTAHGVVVDVLTSCAAVNRRLAHMAFAPRLALYSW----- 371
 QY 378 PARTDGGGGSTPAMASATEANGAEPFA--GIHWISQILYHIGTWLDSETHPLGM 432
 DB 372 -----TGPDQVLKNGLHWISQVILGIGKLLDSELPPLAL 407
 RESULT 14
 ID DH_MOUSE STANDARD: PRT: 396 AA.
 AC 061488;
 DT 15-JUL-1999 (rel. 38, Created)
 DT 15-JUL-1999 (rel. 38, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Desert hedgehog protein precursor (DHH) (HHG-3).
 GN DHH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv;
 RX MEDLINE=94094334; PubMed=7916661;
 RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,
 McMahon J.A., McMahon A.P.;
 "Sonic hedgehog, a member of a family of putative signaling
 molecules, is implicated in the regulation of CNS polarity".
 RT Cell 75:1417-1430(1993).
 RL
 RN
 RP SEQUENCE OF 120-168 FROM N.A.
 RX MEDLINE=95236997; PubMed=7720571;
 RA Zhang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
 Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;

RT "Products, genetic linkage and limb patterning activity of a murine
 RT hedgehog gene."
 RL Development 120:3339-3353(1994).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL. ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A
 CC SPERMATOCYTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES
 CC DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ADULT TESTES. NOT EXPRESSED IN
 CC LIMB BUDS.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 DR EMBL: X76292; CAA53924.1; -.
 DR HSBP: Q62226; LVNH.
 DR MEROPS: C46.004; -.
 DR MKD: MGI:94691; Dh.
 DR InterPro: IPR000320; HH_signal.
 DR InterPro: IPR001767; Hedgehog_hnt.
 DR InterPro: IPR003586; Hedgehog_hntC.
 DR InterPro: IPR003587; Hedgehog_hntn.
 DR InterPro: IPR002203; Intein.
 DR InterPro: IPR001657; SotichH.
 DR Pfam: PF01079; HntC_1.
 DR Pfam: PF01085; HH_signal; 1.
 DR PRINTS: PR00632; SONICHHOG.
 DR ProDom: PD003042; HH_signal; 1.
 DR SMART: SM00305; HntC_1.
 DR SMART: SM00306; Hntn_1.
 DR PROSITE: PS50817; INTEIN_N_TER; 1.
 DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 22
 FT CHAIN 23 396
 FT CHAIN 23 198
 FT CHAIN 199 396
 FT SITE 198 199
 FT SITE 244 244
 FT SITE 268 268
 FT ACT_SITE 271 271
 FT BINDING 198 198
 FT LIPID 23 23
 FT SEQUENCE 396 AA; 43542 MW; AFEE051B590F8 CRC64;
 Query Match 53.3%; Score 1219; DB 1; Length 396;
 Best Local Similarity 58.7%; Pred. No. 1.5e-90;
 Matches 249; Conservative 55; Mismatches 80; Indels 40; Gaps 12;

QY 62 RYEGKITRNSERFKEKLPNTNPDIIFKDEENTGADRLMTORCKDKLNALISVANNQPGV 121
 DB 62 PAECRVTRGSRFPDLYPNVNPDIIFDEENSGADRLMTERCKRVANALIAVANNMPGV 121
 QY 122 RLRTVETGDEDEGHSESLHYEGRAVDITTSDDRSTYGMALRLAVEGDMWYTESKAH 101
 DB 122 RLRTVETGDEDEGHSESLHYEGRAVDITTSDDRSTYGMALRLAVEGDMWYTESKAH 101
 QY 182 IHGCVKANSYAAKSGCPCGSAVHLEGGTKLVKLRGDRYLAADGGRLLYSDFLT 241
 DB 182 IHGCVKANSYAAKSGCPCGSAVHLEGGTKLVKLRGDRYLAADGGRLLYSDFLT 241
 QY 242 FLRDEGAKYFYVETLEPREERILLTAHLLFYAPPHNSGPTPGP--SALFASRYRPG 298
 DB 242 FLRDEGAKYFYVETLEPREERILLTAHLLFYAPPHNSGPTPGP--SALFASRYRPG 298
 QY 299 QRYTVVAERGGDRLLPAVAVSTLREEGANAVPLTAHGTLINRYLASCYAVIEBHW 358
 DB 298 DSVLA--PGGD-ALQPARVARVA-REAVGVFAPLTAHGTLINRYLASCYAVIEBHW 352
 QY 359 AHRAPFRLAHLAALAPARTGSGGSIIPAQSTEARGAEPPTAGIHMYSQLYHIG 418
 DB 359 AHRAPFRLAHLAALAPARTGSGGSIIPAQSTEARGAEPPTAGIHMYSQLYHIG 418
 QY 419 TWLL 422
 DB 392 EELM 395
 RESULT 15
 DHH_HUMAN
 ID DHH_HUMAN STANDARD; PRT; 396 AA.
 AC 043323; 015794;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Desert hedgehog protein precursor (DHH) (HHG-3).
 GN DHH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 ON NCBI_TaxID=9606;
 OX [1]
 RP SEQUENCE FROM N.A.
 RA Tate G., Kishimoto K., Mitsuura T.;
 RT "Expression of Sonic hedgehog and its receptor Patched/Smoothed in
 RT human cancer cell lines and embryonic organs.";
 RL J. Biochem. Mol. Biol. Biophys. 4:27-34(2000).
 RN [2]
 RP SEQUENCE OF 85-178 FROM N.A.
 RC TISSUE=Kidney;
 RA Drummond I.A.;
 RT "Human desert hedgehog.";
 RN Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RP [3]
 RP DISEASE.
 RX PubMed=11017805;
 RA Uehara F., Tate G., Itoh K., Yamaguchi N., Douchi T., Mitsuura T.,
 RA Osame M.;
 RT "A novel mutation of desert hedgehog in a patient with 46,xy partial
 RT gonadal dysgenesis accompanied by multifascicular neuropathy.";
 RL Am. J. Hum. Genet. 67:1302-1305(2000).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL. ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A
 CC SPERMATOCYTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES
 CC DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IT IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN

CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESPECTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING. WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- DISEASE: Defects in DHH may be the cause of partial gonadal
 CC dysgenesis (PGD) with minifascicular neuropathy. PGD is
 CC characterized by the presence of a testis on one side and a streak
 CC or an absent gonad at the other, persistence of Mullerian duct
 CC structures, and a variable degree of genital ambiguity.
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL: AB010994; BAA24866.1; -
 CC DR EMBL: AB010581; BAA24866.1; JOINED.
 CC DR EMBL: AB010993; BAA24866.1; JOINED.
 CC DR EMBL: U59748; AAB03398.1; -
 CC DR HSP: 062226; LVNH.
 CC DR MEROPS: C46.004; -
 CC DR Genev: HGNC:2865; DHH.
 CC DR MIM: 605423; -
 CC DR MIM: 607080; -
 CC DR InterPro: IPR000320; HH_signal.
 CC DR InterPro: IPR001767; Hedgehog_hinc.
 CC DR InterPro: IPR003586; Hedgehog_hinc.
 CC DR InterPro: IPR003587; Hedgehog_hinc.
 CC DR InterPro: IPR001657; SonicHH.
 CC DR Pfam: PF01079; Hhnt; 1.
 CC DR Pfam: PF01085; HH_signal; 1.
 CC DR PRINTS: PR00632; SONICHHOG.
 CC DR ProDom: PD003042; HH_signal; 1.
 CC DR SMART: SM00305; Hhnt; 1.
 CC DR SMART: SM00306; Hhnt; 1.
 CC DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 CC Signal; Lipoprotein; Palmitate.
 CC KW SIGNAL 1 22
 CC FT CHAIN 23 396 DESERT HEDGEHOG PROTEIN.
 CC FT CHAIN 23 198 DESERT HEDGEHOG PROTEIN N-PRODUCT.
 CC FT CHAIN 199 396 DESERT HEDGEHOG PROTEIN C-PRODUCT.
 CC FT SITE 198 199 CLEAVAGE (AUTO-) (BY SIMILARITY).
 CC FT SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY
 CC SIMILARITY).
 CC FT SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY
 CC SIMILARITY).
 CC FT ACT_SITE 271 271 ESSENTIAL FOR AUTO-CLEAVAGE (BY
 CC SIMILARITY).
 CC FT BINDING 198 198 CHOLESTEROL (BY SIMILARITY).
 CC FT LIPID 23 23 PALMITATE (BY SIMILARITY).
 CC FT CONFLICT 177 177 E -> G (IN REF. 2).
 CC SQ SEQUENCE 396 AA; 43577 MW; FCE4FB21972C3AD5 CRC64;
 CC
 CC Query Match 53.2%; Score 1217; DB 1; Length 396;
 CC Best Local Similarity 58.7%; Pred. No. 2.1e-90;
 CC Matches 249; Conservative 53; Mismatches 82; Indels 40; Gaps 12;
 CC
 CC 4 LARCFVLIVASSLVCPLACGPRG-GGKRRHPKK-LTPLAYKQIFPNAVEKTIGASG 61
 CC 7 LPLPLCLALIA-----LPQSCGPGRGPGVGRRIYARKQLVPLLYKQFVGPVETIGASG 61
 CC 62 RYEGKITRNSERKELTPYNPDIJFKDEENTGADRLMTORCKDKLALAIQYNQMPGV 121
 CC 62 PAEGRYARSGSERBRLDVPYNPDIJFKDEENGADRLMTORCKERKNALAIQNMMPGV 121
 CC 122 RLRTVTEGMDGDGHHSESLHYEGRAVDITTSDRRSKYGMLARLAVEAGFDWVYESKAH 181

DB 122 RLRTVTEGMDGDGHHSESLHYEGRAVDITTSDRRSKYGMLARLAVEAGFDWVYESKAH 181
 OY 182 IHOSVAENSVAAKSGCGFPGSATVHLEQGGTFLVDLRPGDVLADDOGRLLYSDFLT 241
 DB 182 VHSYVAADNSLAVRAGCGFPGNATVTLMSGKERKGLNELRGDVLADASGRVPTPVLL 241
 OY 242 FLDRDGAKRVYVETLEPRERLLTAAHLFVAPHNSGPTPGP---SALFASVRPG 298
 DB 242 FLDRDQRRASFVAVETEMPRLTLTPMHLVFAA----RGPARAPGDFAPVARRLRAG 297
 OY 299 QRYTVVAENGDRRLPAVHVSYTLREERAGAVPLTANGTILINRYLASCAVIEESHW 358
 DB 298 DSVLA---PGGD-ALRPARVARVA-REEAVGVFAPLTAHGTILVNDVYLASCYAVLESHOW 352
 OY 359 AHRAPFRLAHLAALAPARTDGGGSIIPAQASATEARCAEPTAGIHWYSOLLYHTG 418
 DB 353 AHRAPFRLAHLA-LEALLP-----GGAV-----QPT-GMHWYSLYLRUA 391
 OY 419 TWLL 422
 DB 392 EELL 395

Search completed: February 20, 2003, 10:12:39
 Job time : 15.409 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:10:28 ; Search time 7.84443 Seconds
(without alignments)
1423.285 Million cell updates/sec

Title: US-09-827-110A-13

Perfect score: 2289 1 MLLILARCLVLILASSLLVC.....GTWILDSETHPIGLMAVKSS 437

Sequence: 1 MLLILARCLVLILASSLLVC.....GTWILDSETHPIGLMAVKSS 437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 2554876 residues 140259

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCr_NEW_PUB pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubppaa/PCrUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2289	100.0	437	8 US-08-900-220C-13	Sequence 13, Appl
2	2289	100.0	437	9 US-08-883-848A-13	Sequence 13, Appl
3	2289	100.0	437	10 US-09-151-999-13	Sequence 13, Appl
4	2286	99.9	437	9 US-09-990-046-14	Sequence 14, Appl
5	2286	99.9	437	9 US-09-969-520A-1	Sequence 1, Appl
6	2286	99.9	437	9 US-10-013-310-1	Sequence 16, Appl
7	2286	99.9	437	9 US-09-733-634-16	Sequence 37, Appl
8	2286	99.9	437	9 US-09-969-520A-3	Sequence 3, Appl
9	2280	99.5	437	9 US-09-969-520A-10	Sequence 10, Appl
10	2277	99.5	437	9 US-09-969-520A-11	Sequence 11, Appl
11	2276	99.4	437	9 US-08-968-520A-2	Sequence 2, Appl
12	2272	99.3	437	9 US-09-969-520A-9	Sequence 4, Appl
13	2270	99.2	437	9 US-09-969-520A-4	Sequence 8, Appl
14	2266	99.0	437	9 US-09-969-520A-6	Sequence 7, Appl
15	2263	98.9	437	9 US-09-969-520A-7	Sequence 5, Appl
16	2254	98.5	437	9 US-09-968-520A-14	Sequence 14, Appl
17	2237	97.7	437	9 US-09-733-634-14	Sequence 15, Appl
18	2065	90.2	475	8 US-08-900-220C-15	Sequence 15, Appl
19	2065	90.2	475	9 US-09-883-848A-15	Sequence 15, Appl
20	2065	90.2	475	10 US-09-021-660A-39	Sequence 39, Appl
21	2065	90.2	475	10 US-09-151-999-15	Sequence 15, Appl
22	2065	90.2	475	8 US-08-900-220C-10	Sequence 10, Appl
23	1788	78.1	425	8 US-08-883-848A-10	Sequence 34, Appl
24	1788	78.1	425	10 US-09-021-660A-34	Sequence 10, Appl
25	1788	78.1	425	10 US-09-151-999-10	Sequence 38, Appl
26	1788	78.1	425	8 US-08-900-220C-14	Sequence 14, Appl
27	1517	66.3	418	8 US-08-900-220C-18	Sequence 18, Appl
28	1517	66.3	418	10 US-09-990-046-29	Sequence 29, Appl
29	1466	64.0	418	8 US-08-900-220C-12	Sequence 12, Appl
30	1466	64.0	411	9 US-08-883-848A-12	Sequence 12, Appl
31	1466	64.0	411	9 US-08-883-848A-16	Sequence 16, Appl
32	1464	64.0	416	8 US-08-900-220C-18	Sequence 18, Appl
33	1464	64.0	416	10 US-09-151-999-18	Sequence 18, Appl
34	1464	64.0	411	9 US-08-900-220C-16	Sequence 16, Appl
35	1276.5	55.8	411	9 US-08-883-848A-16	Sequence 16, Appl
36	1276.5	55.8	411	10 US-09-151-999-16	Sequence 16, Appl
37	1276.5	55.8	411	9 US-09-733-634-28	Sequence 28, Appl
38	1271.5	55.5	449	9 US-09-990-046-29	Sequence 29, Appl
39	1268	55.4	449	9 US-08-900-220C-10	Sequence 30, Appl
40	1268	55.4	411	8 US-08-900-220C-12	Sequence 12, Appl
41	1240	54.2	411	9 US-08-883-848A-12	Sequence 12, Appl
42	1240	54.2	411	10 US-09-151-999-12	Sequence 12, Appl
43	1240	54.2	396	8 US-08-900-220C-11	Sequence 11, Appl
44	1219	53.3	396	9 US-09-990-046-13	Sequence 13, Appl
45	1219	53.3	396	9 US-09-990-046-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-900-220C-13
Sequence 13, Application US/08900220C
Patent No. US20020045206A1

GENERAL INFORMATION:
APPLICANT: Miao, Ningning
Wang, Monica
Mahantappa, Nagesh K.
Jin, Ping

TITLE OF INVENTION: Method of Treating Dopaminergic and
GABA-ergic Disorders

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
STREET: ONE POST OFFICE SQUARE
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Ascii (text)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/900,220C
FILING DATE: 24-JUL-1997
CLASSIFICATION: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: ONV-044.01

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-08-900-220C-13

Query Match

Best Local Similarity 100.0%; Score 2289; DB 8; Length 437;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFRGRRHPKKTLPAYKOFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFRGRRHPKKTLPAYKOFIPNVAEKTIGAS 60
QY 61 GRYEGKITRNSERFKEKLPYNNPDIIFKDEENTGADRLMTORCKDKLNALAI SVNMWPG 120
DB 61 GRYEGKITRNSERFKEKLPYNNPDIIFKDEENTGADRLMTORCKDKLNALAI SVNMWPG 120
QY 121 VRLRTGEMDEDEGHSESLHYEGRAVDITTSDRDSKYGMLARLAEAGFDWYYESKA 180
DB 121 VRLRTGEMDEDEGHSESLHYEGRAVDITTSDRDSKYGMLARLAEAGFDWYYESKA 180
QY 181 HHSVKAENSVAAKSGCGPGSATVHLEGGGTRKLVKDLRPGDRVLAADDOGRLLYSDFL 240
DB 181 HHSVKAENSVAAKSGCGPGSATVHLEGGGTRKLVKDLRPGDRVLAADDOGRLLYSDFL 240
QY 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLTFVAPHNDGPTPGPSALFASRVPRGQR 300
DB 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLTFVAPHNDGPTPGPSALFASRVPRGQR 300
QY 301 VYVVAERGGDRRLPAAVHSVTLREBEAGAVAPLTAHGTILINRVLASCAVAIEESHMAH 360
DB 301 VYVVAERGGDRRLPAAVHSVTLREBEAGAVAPLTAHGTILINRVLASCAVAIEESHMAH 360
QY 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPITAGIHMYSQLYHIGTW 420
DB 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPITAGIHMYSQLYHIGTW 420
QY 421 LLDSETHMPLGMAVKSS 437
DB 421 LLDSETHMPLGMAVKSS 437

```

RESULT 2

US-09-883-848A-13

```

; Sequence 13, Application US/09883848A
; Publication No. US20030022819A1
; GENERAL INFORMATION:
; APPLICANT: Ling, L.
; TITLE OF INVENTION: ANGIOGENESIS-MODULATING COMPOSITIONS AND USES
; FILE REFERENCE: CIBT-P01-119
; CURRENT APPLICATION NUMBER: US/09/883,848A
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/211,919
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-883-848A-13

```

Query Match 100.0%; Score 2289; DB 9; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.2e-197;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFRGRRHPKKTLPAYKOFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFRGRRHPKKTLPAYKOFIPNVAEKTIGAS 60
QY 61 GRYEGKITRNSERFKEKLPYNNPDIIFKDEENTGADRLMTORCKDKLNALAI SVNMWPG 120
DB 61 GRYEGKITRNSERFKEKLPYNNPDIIFKDEENTGADRLMTORCKDKLNALAI SVNMWPG 120

```

```

QY 121 VRLRTGEMDEDEGHSESLHYEGRAVDITTSDRDSKYGMLARLAEAGFDWYYESKA 180
DB 121 VRLRTGEMDEDEGHSESLHYEGRAVDITTSDRDSKYGMLARLAEAGFDWYYESKA 180
QY 181 HHSVKAENSVAAKSGCGPGSATVHLEGGGTRKLVKDLRPGDRVLAADDOGRLLYSDFL 240
DB 181 HHSVKAENSVAAKSGCGPGSATVHLEGGGTRKLVKDLRPGDRVLAADDOGRLLYSDFL 240
QY 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLTFVAPHNDGPTPGPSALFASRVPRGQR 300
DB 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLTFVAPHNDGPTPGPSALFASRVPRGQR 300
QY 301 VYVVAERGGDRRLPAAVHSVTLREBEAGAVAPLTAHGTILINRVLASCAVAIEESHMAH 360
DB 301 VYVVAERGGDRRLPAAVHSVTLREBEAGAVAPLTAHGTILINRVLASCAVAIEESHMAH 360
QY 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPITAGIHMYSQLYHIGTW 420
DB 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPITAGIHMYSQLYHIGTW 420
QY 421 LLDSETHMPLGMAVKSS 437
DB 421 LLDSETHMPLGMAVKSS 437

```

RESULT 3

US-09-151-999-13

```

; Sequence 13, Application US/09151999
; Patent No. US20020151460A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Elizabeth
; TITLE OF INVENTION: REGULATION OF EPITHELIAL TISSUE BY HEDGEHOG-LIKE
; FILE REFERENCE: ONV-031.02
; CURRENT APPLICATION NUMBER: US/09/151,999
; EARLIER FILING DATE: 1998-08-11
; EARLIER APPLICATION NUMBER: 08/955,552
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 437
; TYPE: PRT
; ORGANISM: murine Shh
US-09-151-999-13

```

Query Match 100.0%; Score 2289; DB 10; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.2e-197;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFRGRRHPKKTLPAYKOFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFRGRRHPKKTLPAYKOFIPNVAEKTIGAS 60
QY 61 GRYEGKITRNSERFKEKLPYNNPDIIFKDEENTGADRLMTORCKDKLNALAI SVNMWPG 120
DB 61 GRYEGKITRNSERFKEKLPYNNPDIIFKDEENTGADRLMTORCKDKLNALAI SVNMWPG 120
QY 121 VRLRTGEMDEDEGHSESLHYEGRAVDITTSDRDSKYGMLARLAEAGFDWYYESKA 180
DB 121 VRLRTGEMDEDEGHSESLHYEGRAVDITTSDRDSKYGMLARLAEAGFDWYYESKA 180
QY 181 HHSVKAENSVAAKSGCGPGSATVHLEGGGTRKLVKDLRPGDRVLAADDOGRLLYSDFL 240
DB 181 HHSVKAENSVAAKSGCGPGSATVHLEGGGTRKLVKDLRPGDRVLAADDOGRLLYSDFL 240
QY 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLTFVAPHNDGPTPGPSALFASRVPRGQR 300
DB 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLTFVAPHNDGPTPGPSALFASRVPRGQR 300
QY 301 VYVVAERGGDRRLPAAVHSVTLREBEAGAVAPLTAHGTILINRVLASCAVAIEESHMAH 360
DB 301 VYVVAERGGDRRLPAAVHSVTLREBEAGAVAPLTAHGTILINRVLASCAVAIEESHMAH 360

```

QY 361 RAFAPEFLAALLAALAPARTDGGGSSIPAASATFARGAEPTAGIHWXSOLLHIGTW 420
DB 361 RAFAPEFLAALLAALAPARTDGGGSSIPAASATFARGAEPTAGIHWXSOLLHIGTW 420
QY 421 LLDSETHMPLGMAVKSS 437
DB 421 LLDSETHMPLGMAVKSS 437

RESULT 4

US-09-990-046-14
; Sequence 14, Application US/09990046
; Patent No. US20020156245A1
; GENERAL INFORMATION:
; APPLICANT: de Sauvage, Frederic
; APPLICANT: Carpenter, David A.
; TITLE OF INVENTION: Patched-2
; FILE REFERENCE: P1405R1
; CURRENT APPLICATION NUMBER: US/09/990,046
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/293,505
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-15
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 14
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-990-046-14

Query Match 99.9%; Score 2286; DB 9; Length 437;
Best Local Similarity 99.8%; Pred. No. 2.3e-197;
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLLARCFVLIASSILVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIASSILVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
QY 61 GREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNOMP 120
DB 61 GREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNOMP 120
QY 121 VRLVTEGMEDEGHSESLHYEGRAVDITTSRDSKYGMLARLAVEAGFDWVYESKA 180
DB 121 VRLVTEGMEDEGHSESLHYEGRAVDITTSRDSKYGMLARLAVEAGFDWVYESKA 180
QY 181 HHCISVKAENSVAKSGCCPGSATYHLEGGTKLVKDLRPGDRVLAADOGRLYSDFL 240
DB 181 HHCISVKAENSVAKSGCCPGSATYHLEGGTKLVKDLRPGDRVLAADOGRLYSDFL 240
QY 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLLFVAAPHNDGPTPGPSALFASRVPRGQR 300
DB 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLLFVAAPHNDGPTPGPSALFASRVPRGQR 300
QY 301 VYVVAERGDRRLPAVAHVSVTLREEBAGAYADLTAGTLLINRVLASCAVIEESHMAH 360
DB 301 VYVVAERGDRRLPAVAHVSVTLREEBAGAYADLTAGTLLINRVLASCAVIEESHMAH 360
QY 361 RAFAPEFLAALLAALAPARTDGGGSSIPAASATFARGAEPTAGIHWXSOLLHIGTW 420
DB 361 RAFAPEFLAALLAALAPARTDGGGSSIPAASATFARGAEPTAGIHWXSOLLHIGTW 420
QY 421 LLDSETHMPLGMAVKSS 437
DB 421 LLDSETHMPLGMAVKSS 437

RESULT 5
US-09-969-520A-1
; Sequence 1, Application US/09969520A
; Patent No. US20020177163A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE

; APPLICANT: BEACHY, Philip A.
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHE
; FILE REFERENCE: JHU1670-1
; CURRENT APPLICATION NUMBER: US/09/969,520A
; CURRENT FILING DATE: 2002-06-04
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-969-520A-1

Query Match 99.9%; Score 2286; DB 9; Length 437;
Best Local Similarity 99.8%; Pred. No. 2.3e-197;
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLLARCFVLIASSILVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIASSILVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
QY 61 GREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNOMP 120
DB 61 GREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNOMP 120
QY 121 VRLVTEGMEDEGHSESLHYEGRAVDITTSRDSKYGMLARLAVEAGFDWVYESKA 180
DB 121 VRLVTEGMEDEGHSESLHYEGRAVDITTSRDSKYGMLARLAVEAGFDWVYESKA 180
QY 181 HHCISVKAENSVAKSGCCPGSATYHLEGGTKLVKDLRPGDRVLAADOGRLYSDFL 240
DB 181 HHCISVKAENSVAKSGCCPGSATYHLEGGTKLVKDLRPGDRVLAADOGRLYSDFL 240
QY 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLLFVAAPHNDGPTPGPSALFASRVPRGQR 300
DB 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLLFVAAPHNDGPTPGPSALFASRVPRGQR 300
QY 301 VYVVAERGDRRLPAVAHVSVTLREEBAGAYADLTAGTLLINRVLASCAVIEESHMAH 360
DB 301 VYVVAERGDRRLPAVAHVSVTLREEBAGAYADLTAGTLLINRVLASCAVIEESHMAH 360
QY 361 RAFAPEFLAALLAALAPARTDGGGSSIPAASATFARGAEPTAGIHWXSOLLHIGTW 420
DB 361 RAFAPEFLAALLAALAPARTDGGGSSIPAASATFARGAEPTAGIHWXSOLLHIGTW 420
QY 421 LLDSETHMPLGMAVKSS 437
DB 421 LLDSETHMPLGMAVKSS 437

RESULT 6
US-10-013-310-1
; Sequence 1, Application US/10013310
; Publication No. US20020192216A1
; GENERAL INFORMATION:
; APPLICANT: Lamb, Jonathon Robert
; APPLICANT: Hoyne, Gerard Francis
; APPLICANT: Dallman, Margaret Jane
; TITLE OF INVENTION: Therapeutic Use
; FILE REFERENCE: 674525-2003
; CURRENT APPLICATION NUMBER: US/10/013,310
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: PCT/GB00/02191
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: UK 9913350.6
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: UK 9921953.7
; PRIOR FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1

LENGTH: 437
 TYPE: PRT
 ORGANISM: House Mouse
 US-10-013-310-1

Query Match
 Best Local Similarity 99.9%; Score 2286; DB 9; Length 437;
 Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MLLLLARCFVLIASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
OY 61 GYEGKITRNSERFELPNYNPDIIFKDEENTGADRLMTQCKKLNLAISVNMWPG 120
DB 61 GYEGKITRNSERFELPNYNPDIIFKDEENTGADRLMTQCKKLNLAISVNMWPG 120
OY 121 VRLRTEGMDGDGHSEESLHYEGRAVDITTSDRDSKYGMRLARLAVAGFDMVYTESKA 180
DB 121 VRLRTEGMDGDGHSEESLHYEGRAVDITTSDRDSKYGMRLARLAVAGFDMVYTESKA 180
OY 181 HHCsvKAENSVAAKSGCGPFSATVHLEOGGTRKYKDLRPGDRLAADOGRLLYSDFL 240
DB 181 HHCsvKAENSVAAKSGCGPFSATVHLEOGGTRKYKDLRPGDRLAADOGRLLYSDFL 240
OY 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRYRPGQR 300
DB 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRYRPGQR 300
OY 301 VYVAERGGDRLLPAAVHSTVTLREEGAGAYPLTAHGTTILINVLASCVAVIEESHMAH 360
DB 301 VYVAERGGDRLLPAAVHSTVTLREEGAGAYPLTAHGTTILINVLASCVAVIEESHMAH 360
OY 361 RAFAFRLAHLAALAPARTDGGGSIIPAOSATEARGAEPPIAGIMYSOLLHYIGTW 420
DB 361 RAFAFRLAHLAALAPARTDGGGSIIPAOSATEARGAEPPIAGIMYSOLLHYIGTW 420
OY 421 LLDSETHMPLGMAVKSS 437
DB 421 LLDSETHMPLGMAVKSS 437

```

RESULT 7

US-09-733-634-16
 Sequence 16, Application US/09733634
 Publication No. US20030013646A1
 GENERAL INFORMATION:
 APPLICANT: Massachusetts General Hospital
 TITLE OF INVENTION: Method to stimulate Insulin production by pancreatic b-cells
 FILE REFERENCE: 17633/1240
 CURRENT APPLICATION NUMBER: US/09/733,634
 CURRENT FILING DATE: 2000-12-08
 PRIOR APPLICATION NUMBER: US 60/170,282
 PRIOR FILING DATE: 1999-12-10
 NUMBER OF SEQ ID NOS: 32
 SOFTWARE: Patentn version 3.0
 SEQ ID NO 16
 LENGTH: 437
 TYPE: PRT
 ORGANISM: Mus musculus
 US-09-733-634-16

Query Match
 Best Local Similarity 99.9%; Score 2286; DB 9; Length 437;
 Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MLLLLARCFVLIASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
OY 61 GYEGKITRNSERFELPNYNPDIIFKDEENTGADRLMTQCKKLNLAISVNMWPG 120
DB 61 GYEGKITRNSERFELPNYNPDIIFKDEENTGADRLMTQCKKLNLAISVNMWPG 120
OY 121 VRLRTEGMDGDGHSEESLHYEGRAVDITTSDRDSKYGMRLARLAVAGFDMVYTESKA 180
DB 121 VRLRTEGMDGDGHSEESLHYEGRAVDITTSDRDSKYGMRLARLAVAGFDMVYTESKA 180
OY 181 HHCsvKAENSVAAKSGCGPFSATVHLEOGGTRKYKDLRPGDRLAADOGRLLYSDFL 240
DB 181 HHCsvKAENSVAAKSGCGPFSATVHLEOGGTRKYKDLRPGDRLAADOGRLLYSDFL 240
OY 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRYRPGQR 300
DB 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRYRPGQR 300

```

```

OY 121 VRLRTEGMDGDGHSEESLHYEGRAVDITTSDRDSKYGMRLARLAVAGFDMVYTESKA 180
DB 121 VRLRTEGMDGDGHSEESLHYEGRAVDITTSDRDSKYGMRLARLAVAGFDMVYTESKA 180
OY 181 HHCsvKAENSVAAKSGCGPFSATVHLEOGGTRKYKDLRPGDRLAADOGRLLYSDFL 240
DB 181 HHCsvKAENSVAAKSGCGPFSATVHLEOGGTRKYKDLRPGDRLAADOGRLLYSDFL 240
OY 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRYRPGQR 300
DB 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRYRPGQR 300
OY 301 VYVAERGGDRLLPAAVHSTVTLREEGAGAYPLTAHGTTILINVLASCVAVIEESHMAH 360
DB 301 VYVAERGGDRLLPAAVHSTVTLREEGAGAYPLTAHGTTILINVLASCVAVIEESHMAH 360
OY 361 RAFAFRLAHLAALAPARTDGGGSIIPAOSATEARGAEPPIAGIMYSOLLHYIGTW 420
DB 361 RAFAFRLAHLAALAPARTDGGGSIIPAOSATEARGAEPPIAGIMYSOLLHYIGTW 420
OY 421 LLDSETHMPLGMAVKSS 437
DB 421 LLDSETHMPLGMAVKSS 437

```

RESULT 8

US-09-021-660A-37
 Sequence 37, Application US/09021660A
 Patent No. US2001004168A1
 GENERAL INFORMATION:
 APPLICANT: Baron, M.
 APPLICANT: Balausoff, M.
 TITLE OF INVENTION: METHODS FOR MODULATING HEMATOPOIESIS AND VASCULAR
 TITLE OF INVENTION: GROWTH
 FILE REFERENCE: HUIP-P01-060
 CURRENT APPLICATION NUMBER: US/09/021,660A
 CURRENT FILING DATE: 2001-08-27
 PRIOR APPLICATION NUMBER: 60/037,513
 PRIOR FILING DATE: 1997-02-10
 PRIOR APPLICATION NUMBER: 60/049,763
 PRIOR FILING DATE: 1997-06-16
 NUMBER OF SEQ ID NOS: 42
 SOFTWARE: Patentn Ver. 2.1
 SEQ ID NO 37
 LENGTH: 437
 TYPE: PRT
 ORGANISM: Mus musculus
 US-09-021-660A-37

Query Match
 Best Local Similarity 99.9%; Score 2286; DB 10; Length 437;
 Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MLLLLARCFVLIASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
OY 61 GYEGKITRNSERFELPNYNPDIIFKDEENTGADRLMTQCKKLNLAISVNMWPG 120
DB 61 GYEGKITRNSERFELPNYNPDIIFKDEENTGADRLMTQCKKLNLAISVNMWPG 120
OY 121 VRLRTEGMDGDGHSEESLHYEGRAVDITTSDRDSKYGMRLARLAVAGFDMVYTESKA 180
DB 121 VRLRTEGMDGDGHSEESLHYEGRAVDITTSDRDSKYGMRLARLAVAGFDMVYTESKA 180
OY 181 HHCsvKAENSVAAKSGCGPFSATVHLEOGGTRKYKDLRPGDRLAADOGRLLYSDFL 240
DB 181 HHCsvKAENSVAAKSGCGPFSATVHLEOGGTRKYKDLRPGDRLAADOGRLLYSDFL 240
OY 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRYRPGQR 300
DB 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRYRPGQR 300

```



```

Db 241 TFLDRBEGAKKVFYVETLEPRERLLTAAHLFVAPRHNDSGTPGPSALFASRVPRGQR 300
Oy 301 VYVAERGGDRRLPAVAHSVTLREBEAGAYABLTAGTLLINRVLASCAVIEESHMAH 360
Db 301 VYVAERGGDRRLPAVAHSVTLREBEAGAYABLTAGTLLINRVLASCAVIEESHMAH 360
Oy 361 RAFAPRRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPAGIHWYSOLLYHIGTW 420
Db 361 RAFAPRRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPAGIHWYSOLLYHIGTW 420
Oy 421 LLDSETHMPLGMAVKSS 437
Db 421 LLDSETHMPLGMAVKSS 437

RESULT 9
US-09-969-520A-3
; Sequence 3, Application US/09969520A
; Patent No. US20020177163A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: BEACHY, Phillip A.
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED
; FILE REFERENCE: JHU1670-1
; CURRENT APPLICATION NUMBER: US/09/969,520A
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/235,153
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 437
; TYPE: PRP
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Altered sonic hedgehog protein
US-09-969-520A-3

Query Match 99.6%; Score 2280; DB 9; Length 437;
Best local Similarity 99.5%; Pred. No. 7,9e-197;
Matches 435; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MLLLARCFLYIASSLLVCPGLACGPGRGFGRRHPKKTPLATYKQFIPNVAEKTIGAS 60
Db 1 MLLLARCFLYIASSLLVCPGLACGPGRGFGRRHPKKTPLATYKQFIPNVAEKTIGAS 60
Oy 61 GREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALATISVMNOMP 120
Db 61 GREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALATISVMNOMP 120
Oy 121 VRLVTEGMDGDHNSHESLHYEGRAVDITTSDBRSKYGMLARLAVEAGFDWVYESKA 180
Db 121 VRLVTEGMDGDHNSHESLHYEGRAVDITTSDBRSKYGMLARLAVEAGFDWVYESKA 180
Oy 181 HHCYSKAENSVAAGSGCGPGSATYHLEGGGTKLYKDLPRGDRVLAADOGRLYSDFL 240
Db 181 HHCYSKAENSVAAGSGCGPGSATYHLEGGGTKLYKDLPRGDRVLAADOGRLYSDFL 240
Oy 241 TFLDRBEGAKKVFYVETLEPRERLLTAAHLFVAPRHNDSGTPGPSALFASRVPRGQR 300
Db 241 TFLDRBEGAKKVFYVETLEPRERLLTAAHLFVAPRHNDSGTPGPSALFASRVPRGQR 300
Oy 301 VYVAERGGDRRLPAVAHSVTLREBEAGAYABLTAGTLLINRVLASCAVIEESHMAH 360
Db 301 VYVAERGGDRRLPAVAHSVTLREBEAGAYABLTAGTLLINRVLASCAVIEESHMAH 360
Oy 361 RAFAPRRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPAGIHWYSOLLYHIGTW 420
Db 361 RAFAPRRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPAGIHWYSOLLYHIGTW 420
Oy 421 LLDSETHMPLGMAVKSS 437
Db 421 LLDSETHMPLGMAVKSS 437

```

```

RESULT 10
US-09-969-520A-10
; Sequence 10, Application US/09969520A
; Patent No. US20020177163A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: BEACHY, Phillip A.
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHE
; FILE REFERENCE: JHU1670-1
; CURRENT APPLICATION NUMBER: US/09/969,520A
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/235,153
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 437
; TYPE: PRP
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Altered sonic hedgehog protein
US-09-969-520A-10

Query Match 99.5%; Score 2277; DB 9; Length 437;
Best local Similarity 99.3%; Pred. No. 1,5e-196;
Matches 434; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MLLLARCFLYIASSLLVCPGLACGPGRGFGRRHPKKTPLATYKQFIPNVAEKTIGAS 60
Db 1 MLLLARCFLYIASSLLVCPGLACGPGRGFGRRHPKKTPLATYKQFIPNVAEKTIGAS 60
Oy 61 GREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALATISVMNOMP 120
Db 61 GREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALATISVMNOMP 120
Oy 121 VRLVTEGMDGDHNSHESLHYEGRAVDITTSDBRSKYGMLARLAVEAGFDWVYESKA 180
Db 121 VRLVTEGMDGDHNSHESLHYEGRAVDITTSDBRAKYGMLARLAVEAGFDWVYESKA 180
Oy 181 HHCYSKAENSVAAGSGCGPGSATYHLEGGGTKLYKDLPRGDRVLAADOGRLYSDFL 240
Db 181 HHCYSKAENSVAAGSGCGPGSATYHLEGGGTKLYKDLPRGDRVLAADOGRLYSDFL 240
Oy 241 TFLDRBEGAKKVFYVETLEPRERLLTAAHLFVAPRHNDSGTPGPSALFASRVPRGQR 300
Db 241 TFLDRBEGAKKVFYVETLEPRERLLTAAHLFVAPRHNDSGTPGPSALFASRVPRGQR 300
Oy 301 VYVAERGGDRRLPAVAHSVTLREBEAGAYABLTAGTLLINRVLASCAVIEESHMAH 360
Db 301 VYVAERGGDRRLPAVAHSVTLREBEAGAYABLTAGTLLINRVLASCAVIEESHMAH 360
Oy 361 RAFAPRRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPAGIHWYSOLLYHIGTW 420
Db 361 RAFAPRRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPAGIHWYSOLLYHIGTW 420
Oy 421 LLDSETHMPLGMAVKSS 437
Db 421 LLDSETHMPLGMAVKSS 437

RESULT 11
US-09-969-520A-11
; Sequence 11, Application US/09969520A
; Patent No. US20020177163A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: BEACHY, Phillip A.
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHE
; FILE REFERENCE: JHU1670-1
; CURRENT APPLICATION NUMBER: US/09/969,520A
; CURRENT FILING DATE: 2002-06-04

```

;; PRIOR APPLICATION NUMBER: US 60/235,153
;; PRIOR FILING DATE: 2000-09-22
;; NUMBER OF SEQ ID NOS: 17
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 11
;; LENGTH: 437
;; TYPE: PRT
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Altered sonic hedgehog protein
US-09-969-520A-11

Query Match 99.5%; Score 2272; DB 9; Length 437;
Best Local Similarity 99.3%; Pred. No. 1,5e-196;
Matches 434; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLLIARCFVLIASSLLVCPGLACGPGRGFGKRRHPKRLTPLAYKQFIPNVAEKTIGAS 60
Db 1 MLLIARCFVLIASSLLVCPGLACGPGRGFGKRRHPKRLTPLAYKQFIPNVAEKTIGAS 60
QY 61 GREGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALAI SVMQMPG 120
Db 61 GREGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALAI SVMQMPG 120
QY 121 VRLVTEGMDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVAGFDWVYYESKA 180
Db 121 VRLVTEGMDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVAGFDWVYYEAAA 180
QY 181 HHCYKAENSVAKSGCGPGSATVHLDEGGTKLVKDLRPGDRLVLAADOGRLYSDFL 240
Db 181 HHCYKAENSVAKSGCGPGSATVHLDEGGTKLVKDLRPGDRLVLAADOGRLYSDFL 240
QY 241 TELDRDEGAKKVFYETLEPRERLLTAAHLFVAPHNDSGPTPGPSALFASRVPRGOR 300
Db 241 TELDRDEGAKKVFYETLEPRERLLTAAHLFVAPHNDSGPTPGPSALFASRVPRGOR 300
QY 301 VYVAERGGDRLLPAVHSVTLREEAGAYAPLTAHGTLLINRYLASCYAVEEHSMAH 360
Db 301 VYVAERGGDRLLPAVHSVTLREEAGAYAPLTAHGTLLINRYLASCYAVEEHSMAH 360
QY 361 RAFAFRLHALLAALAPARTDGGGGSIPAAQSATEARGAEPYAGIHWSQLYHIGTW 420
Db 361 RAFAFRLHALLAALAPARTDGGGGSIPAAQSATEARGAEPYAGIHWSQLYHIGTW 420
QY 421 LLDSETHMPLGMAVKSS 437
Db 421 LLDSETHMPLGMAVKSS 437

RESULT 12

US-09-969-520A-2
;; Sequence 2, Application US/09969520A
;; Patent No. US20020177163A1
;; GENERAL INFORMATION:
;; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
;; APPLICANT: BEACHY, Philip A.
;; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED
;; FILE REFERENCE: JH01670-1
;; CURRENT APPLICATION NUMBER: US/09/969, 520A
;; PRIOR FILING DATE: 2002-06-04
;; PRIOR APPLICATION NUMBER: US 60/235,153
;; NUMBER OF SEQ ID NOS: 17
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 2
;; LENGTH: 437
;; TYPE: PRT
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Altered sonic hedgehog protein
US-09-969-520A-2

Query Match 99.4%; Score 2276; DB 9; Length 437;

Best Local Similarity 99.5%; Pred. No. 1.8e-196;
Matches 435; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLLIARCFVLIASSLLVCPGLACGPGRGFGKRRHPKRLTPLAYKQFIPNVAEKTIGAS 60
Db 1 MLLIARCFVLIASSLLVCPGLACGPGRGFGKRRHPKRLTPLAYKQFIPNVAEKTIGAS 60
QY 61 GREGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALAI SVMQMPG 120
Db 61 GREGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALAI SVMQMPG 120
QY 121 VRLVTEGMDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVAGFDWVYYESKA 180
Db 121 VRLVTEGMDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVAGFDWVYYESKA 180
QY 181 HHCYKAENSVAKSGCGPGSATVHLDEGGTKLVKDLRPGDRLVLAADOGRLYSDFL 240
Db 181 HHCYKAENSVAKSGCGPGSATVHLDEGGTKLVKDLRPGDRLVLAADOGRLYSDFL 240
QY 241 TELDRDEGAKKVFYETLEPRERLLTAAHLFVAPHNDSGPTPGPSALFASRVPRGOR 300
Db 241 TELDRDEGAKKVFYETLEPRERLLTAAHLFVAPHNDSGPTPGPSALFASRVPRGOR 300
QY 301 VYVAERGGDRLLPAVHSVTLREEAGAYAPLTAHGTLLINRYLASCYAVEEHSMAH 360
Db 301 VYVAERGGDRLLPAVHSVTLREEAGAYAPLTAHGTLLINRYLASCYAVEEHSMAH 360
QY 361 RAFAFRLHALLAALAPARTDGGGGSIPAAQSATEARGAEPYAGIHWSQLYHIGTW 420
Db 361 RAFAFRLHALLAALAPARTDGGGGSIPAAQSATEARGAEPYAGIHWSQLYHIGTW 420
QY 421 LLDSETHMPLGMAVKSS 437
Db 421 LLDSETHMPLGMAVKSS 437

RESULT 13

US-09-969-520A-9
;; Sequence 9, Application US/09969520A
;; Patent No. US20020177163A1
;; GENERAL INFORMATION:
;; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
;; APPLICANT: BEACHY, Philip A.
;; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHE
;; FILE REFERENCE: JH01670-1
;; CURRENT APPLICATION NUMBER: US/09/969, 520A
;; PRIOR FILING DATE: 2002-06-04
;; PRIOR APPLICATION NUMBER: US 60/235,153
;; NUMBER OF SEQ ID NOS: 17
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 9
;; LENGTH: 437
;; TYPE: PRT
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Altered sonic hedgehog protein
US-09-969-520A-9

Query Match 99.3%; Score 2272; DB 9; Length 437;
Best Local Similarity 99.3%; Pred. No. 4.1e-196;
Matches 434; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLLIARCFVLIASSLLVCPGLACGPGRGFGKRRHPKRLTPLAYKQFIPNVAEKTIGAS 60
Db 1 MLLIARCFVLIASSLLVCPGLACGPGRGFGKRRHPKRLTPLAYKQFIPNVAEKTIGAS 60
QY 61 GREGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALAI SVMQMPG 120
Db 61 GREGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALAI SVMQMPG 120
QY 121 VRLVTEGMDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVAGFDWVYYESKA 180
Db 121 VRLVTEGMDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVAGFDWVYYESKA 180

Db 121 VKLVTGMDGHDGHSSESLHYEGRAVDITTSDBRSKYGMALAVAEAGFDWYYESKA 180
QY 181 HHCSSVAENSVAASGCGCFPGSATVHLEOGGKTLVLDLRPGDVLAAADGGRLYSDFL 240
Db 181 HHCSSVAENSVAASGCGCFPGSATVHLEOGGKTLVLDLRPGDVLAAADGGRLYSDFL 240
QY 241 TFLDRDGAKKVFYVIELEPERRELLTAAHLLEVAAPHNDGPTPGPSALFASVRPGQR 300
Db 241 TFLDRDGAKKVFYVIELEPERRELLTAAHLLEVAAPHNDGPTPGPSALFASVRPGQR 300
QY 301 VYVAERGGDRLLPAAVHSTLREERAGAYAPLTAGTILINNVLASCAVIEHSHMAH 360
Db 301 VYVAERGGDRLLPAAVHSTLREERAGAYAPLTAGTILINNVLASCAVIEHSHMAH 360
QY 361 RAFAFRLAHLAALAPARTDGGGGSIPAAQSATERGAEPAGIHWYSQLLYHIGTW 420
Db 361 RAFAFRLAHLAALAPARTDGGGGSIPAAQSATERGAEPAGIHWYSQLLYHIGTW 420
QY 421 LLDSETHMPLGMAVKSS 437
Db 421 LLDSETHMPLGMAVKSS 437

RESULT 14

US-09-969-520A-4
; Sequence 4, Application US/09969520A
; Patent No. US20020177163A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: BEACHY, Philip A.
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED
; FILE REFERENCE: JH01670-1
; CURRENT APPLICATION NUMBER: US/09/969,520A
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/235,153
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Altered sonic hedgehog protein
US-09-969-520A-4

Query Match 99.2% Score 2270; DB 9; Length 437;
Best Local Similarity 99.3% Pred. No. 6.2e-195;

Matches 434; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLLLLARCFVILASLLVCPGLACGPGRGFKRRHKPLTPLAYKOFIPVAEKTIGAS 60
Db 1 MLLLLARCFVILASLLVCPGLACGPGRGFKRRHKPLTPLAYKOFIPVAEKTIGAS 60
QY 61 GRYEGKITRNSERKELTPNPNPDIFKDEENTGADRLMTORCKDKLNALASVNMOWPG 120
Db 61 GRYEGKITRNSERKELTPNPNPDIFKDEENTGADRLMTORCKDKLNALASVNMOWPG 120
QY 121 VRLAVTEGMDGHDGHSSESLHYEGRAVDITTSDBRSKYGMALAVAEAGFDWYYESKA 180
Db 121 VRLAVTEGMDGHDGHSSESLHYEGRAVDITTSDBRSKYGMALAVAEAGFDWYYESKA 180
QY 181 HHCSSVAENSVAASGCGCFPGSATVHLEOGGKTLVLDLRPGDVLAAADGGRLYSDFL 240
Db 181 HHCSSVAENSVAASGCGCFPGSATVHLEOGGKTLVLDLRPGDVLAAADGGRLYSDFL 240
QY 241 TFLDRDGAKKVFYVIELEPERRELLTAAHLLEVAAPHNDGPTPGPSALFASVRPGQR 300
Db 241 TFLDRDGAKKVFYVIELEPERRELLTAAHLLEVAAPHNDGPTPGPSALFASVRPGQR 300
QY 301 VYVAERGGDRLLPAAVHSTLREERAGAYAPLTAGTILINNVLASCAVIEHSHMAH 360
Db 301 VYVAERGGDRLLPAAVHSTLREERAGAYAPLTAGTILINNVLASCAVIEHSHMAH 360

QY 361 RAFAFRLAHLAALAPARTDGGGGSIPAAQSATERGAEPAGIHWYSQLLYHIGTW 420
Db 361 RAFAFRLAHLAALAPARTDGGGGSIPAAQSATERGAEPAGIHWYSQLLYHIGTW 420
QY 421 LLDSETHMPLGMAVKSS 437
Db 421 LLDSETHMPLGMAVKSS 437

RESULT 15

US-09-969-520A-8
; Sequence 8, Application US/09969520A
; Patent No. US20020177163A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: BEACHY, Philip A.
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED
; FILE REFERENCE: JH01670-1
; CURRENT APPLICATION NUMBER: US/09/969,520A
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/235,153
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 8
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Altered sonic hedgehog protein
US-09-969-520A-8

Query Match 99.0% Score 2266; DB 9; Length 437;
Best Local Similarity 99.1% Pred. No. 1.4e-195;
Matches 433; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLLLLARCFVILASLLVCPGLACGPGRGFKRRHKPLTPLAYKOFIPVAEKTIGAS 60
Db 1 MLLLLARCFVILASLLVCPGLACGPGRGFKRRHKPLTPLAYKOFIPVAEKTIGAS 60
QY 61 GRYEGKITRNSERKELTPNPNPDIFKDEENTGADRLMTORCKDKLNALASVNMOWPG 120
Db 61 GRYEGKITRNSERKELTPNPNPDIFKDEENTGADRLMTORCKDKLNALASVNMOWPG 120
QY 121 VRLAVTEGMDGHDGHSSESLHYEGRAVDITTSDBRSKYGMALAVAEAGFDWYYESKA 180
Db 121 VRLAVTEGMDGHDGHSSESLHYEGRAVDITTSDBRSKYGMALAVAEAGFDWYYESKA 180
QY 181 HHCSSVAENSVAASGCGCFPGSATVHLEOGGKTLVLDLRPGDVLAAADGGRLYSDFL 240
Db 181 HHCSSVAENSVAASGCGCFPGSATVHLEOGGKTLVLDLRPGDVLAAADGGRLYSDFL 240
QY 241 TFLDRDGAKKVFYVIELEPERRELLTAAHLLEVAAPHNDGPTPGPSALFASVRPGQR 300
Db 241 TFLDRDGAKKVFYVIELEPERRELLTAAHLLEVAAPHNDGPTPGPSALFASVRPGQR 300
QY 301 VYVAERGGDRLLPAAVHSTLREERAGAYAPLTAGTILINNVLASCAVIEHSHMAH 360
Db 301 VYVAERGGDRLLPAAVHSTLREERAGAYAPLTAGTILINNVLASCAVIEHSHMAH 360
QY 361 RAFAFRLAHLAALAPARTDGGGGSIPAAQSATERGAEPAGIHWYSQLLYHIGTW 420
Db 361 RAFAFRLAHLAALAPARTDGGGGSIPAAQSATERGAEPAGIHWYSQLLYHIGTW 420
QY 421 LLDSETHMPLGMAVKSS 437
Db 421 LLDSETHMPLGMAVKSS 437

Search completed: February 20, 2003, 10:14:42
Job time : 8.84443 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:06:28 ; Search time 28.1092 Seconds
(without alignments)
3203.314 Million cell updates/sec

Title: US-09-827-110a-13
Perfect score: 2289
Sequence: 1 MLLILARCLVLIASSLVC.....GTWLDSETHMPLGMVKSS 437

Scoring table: BLAST62
Gap 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriophage:*
17: sp_archaeophages:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1537.5	67.2	434	13	057404 pleurodeles
2	1451	63.4	414	13	09W709 parallelthly
3	1326	57.9	406	13	057567 notophthalm
4	1132.5	49.5	415	5	09U526 branchiosteo
5	1124.5	49.1	415	5	017499 branchiosteo
6	1063.5	46.5	442	13	073803 fugu rubrip
7	1040.5	45.0	447	5	09GRA8 O9gras gryllus him
8	915.5	40.0	410	5	061676 lytechinus
9	842	36.8	161	11	09R179 O9r179 rattus norv
10	736	32.2	150	13	09YGV7 O9ygv7 ambystoma m
11	732	32.0	139	6	09XSI6 O9xsi6 bos taurus
12	710	31.0	138	13	09W6C1 O9w6c1 eleutheroda
13	698	30.5	177	11	09W6Z9 O9w6z9 rattus norv
14	628	27.4	185	5	096699 junonia coe
15	528	23.1	99	13	08G6N4 O8g6n4 scyllorhinu
16	524	22.9	119	13	042128 oryzias lat

17	518	22.6	129	11	09WUP6 O9wup6 rattus norv
18	489	21.4	137	13	042234 O42234 coturnix co
19	395	17.3	87	5	09TX30 O9tx30 anopheles g
20	389	17.0	80	13	042441 O42441 oryzias lat
21	339	14.8	63	13	09T8T4 O9t8t4 rana catesb
22	329	14.4	64	13	09PRF5 O9prf5 oryzias lat
23	295	12.9	60	5	09S2D8 O9s2d8 hydra atten
24	289	12.6	56	13	09O2R0 O9o2r0 oryzias lat
25	234	10.2	49	5	09TX33 O9tx33 hludo medl
26	232	10.1	49	5	09TX31 O9tx31 trilobium c
27	187.5	8.2	125	11	09ESH3 O9esh3 rattus norv
28	175	7.6	557	5	094410 O94410 caenorhabd1
29	175	7.6	557	5	094129 O94129 caenorhabd1
30	174	7.6	54	13	042233 O42233 coturnix co
31	172.5	7.5	48	5	09TX32 O9tx32 strongyloce
32	170	7.4	615	5	P91573 P91573 caenorhabd1
33	153.5	6.7	1207	5	021535 021535 caenorhabd1
34	148	6.5	790	5	022872 022872 caenorhabd1
35	148	6.5	1203	5	021835 021835 caenorhabd1
36	146	6.4	481	5	045992 045992 caenorhabd1
37	140.5	6.1	550	5	045273 045273 caenorhabd1
38	140.5	6.1	550	5	094130 094130 caenorhabd1
39	130.5	5.7	868	5	09XV14 09xv14 caenorhabd1
40	128	5.6	1021	5	09XUV2 09xuv2 caenorhabd1
41	127	5.5	485	5	094128 094128 caenorhabd1
42	118	5.2	598	2	083032 083032 streptomyce
43	117	5.1	1407	16	069826 069826 streptomyce
44	113	4.9	599	16	09L010 09l010 streptomyce
45	106	4.6	969	16	092RRO 092rro rhizobium m

ALIGNMENTS

RESULT 1

057404 PRELIMINARY; PRT; 434 AA.
ID 057404
AC 057404;
DT 01-JUN-1998 (TREMBLERL. 06, Created)
DR 01-JUN-1998 (TREMBLERL. 06, Last sequence update)
DE 01-MAR-2002 (TREMBLERL. 20, Last annotation update)
DE Sonic hedgehog-related protein.
GN pw-SH.
OS Pleurodeles waltl11 (Iberian ribbed newt).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Caudata; Salamandridae; Salamandridae;
OC Pleurodeles.
NCBI_TaxID=8319;
RN [1]
RP SEQUENCE FROM N.A.
RA Caubit X., Nicolas S., Le Parco Y.;
RT "Pleurodeles sonic hedgehog".
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003532; AAB94412.1; -.
DR HSSP; Q62226; IVH.
DR MEROPS; C46.002; -.
DR InterPro: IPR001767; Hedgehog_hnt.
DR InterPro: IPR003586; Hedgehog_hntC.
DR InterPro: IPR003587; Hedgehog_hntN.
DR InterPro: IPR003203; Intein.
DR InterPro: IPR002203; Intein.
DR Pfam: PF01085; HH_signal; 1.
DR Pfam: PF01079; Hnt; 1.
DR PRINTS; PR00632; SONICHHOG.
DR PRODOM; PD003042; HH_signal; 1.
DR SMART; SM00305; HntC; 1.
DR SMART; SM00306; HntN; 1.
SQ SEQUENCE 434 AA; 48421 MW; A9495E367151AE74 CRC64;

Query Match 67.2%; Score 1537.5; DB 13; Length 434;
Best Local Similarity 69.5%; Pred. No. 1.8e-115;
Matches 310; Conservative 38; Mismatches 73; Indels 25; Gaps 6;

OY	2	LLLLACPLVLIASLLACPLACGPGGFCRRHPPKPLPLAYKOPLPNVAEKTGLASG	61
	:	: :	
Dd	4	MILLRDLGGFISTILVPLGCGGPGRGIOQRKRPOKLPLAYKOPLPNVAEKTGLASG	63
OY	62	RYEKGITRNSERFEKELPPNVPDIIIFKDEENTGADRLMTORCKDKLALAI SVANOMPgV	121
	:		
Dd	64	RYEVKITRNSERFOELPPNVPDIIIFKVEENTGADRLMTORCKDKLNSIAI SVANOMPgV	123
OY	122	RLRYTEGDEGCHHSEESLHYEGRAVDITTSDDPRSKYGMILARLAVENGDMWYEEKAH	181
	:		
Dd	124	KLRYTEGDEGCHHSEESLHYEGRAVDITTSDDPRSKYGMILARLAVENGDMWYEEKAH	183
OY	182	IHCYSKAKNSVAARKSGGCFPCSAFVHLEOGCTKVLKDLRPEDRYLADDDGRLTYSDFLT	241
	:		
Dd	184	IHCYSKAKNSVAARKSGGCFPCSAFVHLEOGCTKVLKDLRPEDRYLADDDGRLTYSDFLT	243
OY	242	FLUDRECAKKFYIYIETLEPBERLLTPAHLLEVAAPHNDGSPYGP-----SALF	291
	:		
Dd	244	FMDEETLRKVFYIYIETSLPBERLLTPAHLLEVAAPHNDGSPYGP-----SALF	293
OY	292	ASRYAPGGRVYVVAERGGDRLLTPAAHVSYLREEGGAVALPLAHGHTIILINRYLASCYA	351
	:		
Dd	304	ASSVAPGRV- LTEDREG- KGLREATYDRLYL- DEANGATAVPYAHGTIVIDRYLASCYA	360
OY	352	VIEEHSMAHFAFEPRLAHALAHALAPARTDGGGGGSIIPAQSAATERGAEPTAGIHWMS	411
	:		
Dd	361	VIEEHSMAHFAFAPLRVGGFILTSPFSDYSMS-----PPASQAE-----GVHWMS	408
OY	412	QLLYHICWLLDSEFMHPLGAAVSS	437
	:		
Dd	409	EILRYIGTWVLAADTIHPLGMAASS	434

RESULT 2			
09W709	PRELIMINARY;	PRT;	414 AA.
ID 09W709			
AC 09W709;			
DT 01-NOV-1999 (TREMBREL, 12, Created)			
DT 01-NOV-1999 (TREMBREL, 12, Last sequence update)			
DT 01-MAR-2002 (TREMBREL, 20, Last annotation update)			
DE Sonic hedgehog.			
CN SHH.			
OS Paralleichthys olivaceus (Flounder).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;			
OC Pleuronectidae; Paralleichthyidae; Paralleichthys.			
OX NCBI_Taxid=8255;			
RA [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE-99238226; PubMed-10223710;			
RA Suzuki T., Ichiro O., Kurokawa T.;			
"Retinoic acid given at late embryonic stage depresses sonic hedgehog			
RT and Hoxd-4 expression in the pharyngeal area and induces skeletal			
RT malformation in flounder (Paralleichthys olivaceus) embryos.";			
RL Dev. Growth Differ. 4:143-152(1999).			
DR EMBL: AB029748; BAA82360.1; -			
DR HSSP: Q62226; 1VHN.			
DR MEMOPS: C46_001; -			
DR InterPro: IPR001767; Hedgehog_hint.			
DR InterPro: IPR003586; Hedgehog_hintc.			
DR InterPro: IPR003587; Hedgehog_hintN.			
DR InterPro: IPR003020; HH_signal.			
DR InterPro: IPR002203; Intein.			
DR InterPro: IPR001657; SonicHH.			
DR Pfam: PF01085; HH_signal; 1.			
DR Pfam: PF01079; Hint; 1.			
DR PRINTS: PR00632; SONICHHOG.			
DR ProDom: PD003042; HH_signal; 1.			
DR SMART: SM00305; Hintc; 1.			
DR SMART: SM00306; HintN; 1.			
DR SEQUENCE 414 AA; 45945 MW; 50607BF3DB7CDAA3 CRC64;			

Query Match	63.48;	Score 1451;	DB 13;	Length 414;
Best Local Similarity	65.88;	Pred. No. 1.5e-108;		
Matches 288;	Conservative 53;	Mismatches 71;	Indels 26;	Gaps 7;

```

0Y 2 LLLARCLVLIASSLLVYCPGLACGPGHGFERRRHPKLLPLAYKOTIPNVAEKTLCASC 61
Db 1 MLLMTRIVLAGVLCILSVSSMGCGPGGYGRRRHPKLLPLAYKOTIPNVAEKTLCASC 60
0Y 62 RYEGKITRNSERFKELTPNPNPDIIFKDEENTGADRLMTCORCKDLAALAYSVMQMPGV 121
Db 61 RYEGKITRNSERFKELTPNPNMTDIIFKDEENTGADRLMTCORCKDLAALAYSVMQMPGV 120
0Y 122 RLRTVEGDEDEGHHSEELHTEGRAVDITTSDDRSKYGMILARLAVAGEFDWYYESKAH 181
Db 121 KLRTVEGDEDEGHHSEELHTEGRAVDITTSDDRSKYGLTSRLAVAGEFDWYYESKAH 180
0Y 182 IHCSTKAENSVAAKSGCGFPCSAIVHLEOGSTLVKDLRPEDRYLADDOGRLLYSDFLT 241
Db 181 IHCSTKAENSVAAKSGCGFPCSSVTYTLQDGKTKKVKALQTSDRYLAADHAGOPYTDFIM 240
0Y 242 FLDRDEGAKKVFYIETLEPRERLLTAHLLFYAPHNDSGPT-PPGSALFASHVRPQR 300
Db 241 FLDDSTTRRLFYUIET-DSGOKITTLTAHLLFVGHNSINSTERAHGMSAVFASQVRRPQT 299
0Y 301 YVVV-AEENGGRBLLPRAVHVTLREBEAGYADULTAHTILLIRVLAASYALEESMA 355
Db 300 VFVLDAE-----RLQPVYVKRIYQEH-GSFAVLTAGTVVYDVLASCAVYQDHELA 353
0Y 360 HRAFAFPLALAHALLAALAPARTDGGGGSLPAAOSATAEARGAEPTTAGIHHYSOLLHYGT 419
Db 354 HMAALPARYALAHVYSLLE-----SSQQAASAGND-----GVHHYSKLVQLGT 396
0Y 420 WLLDSETHHPRGMAVKSS 437
Db 397 WLLDHSHTHPLGMSVYPS 414

```

RESULT 3	057567	PRELIMINARY;	PRT;	406 AA.
ID	057567			
AC	057567;			
DT	01-JUN-1998 (TREMBLrel. 06, Created)			
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Hedgehog, segment polarity homolog.			
OS	Trichoplalmus viridescens (Easter n ewt) (Triturus viridescens)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; I			
CC	Amphibia; Batrachia; Caudata; Salamandroides; Salam			
CC	Nonophthalmus.			
OX	NCBI_TaxID=8316;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Stark D.R., Gates P.B., Brookes J.P., Ferretti P.;			
RT	"Hedgehog homologue from <i>Trichoplalmus viridescens</i> ."			
RL	Dev. Dyn. 0:0-0(1998).			
CC	-1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE			
CC	PHOSPHORIBOSYLTRANSFERASE FAMILY.			
DR	EMBL; AF047466; AAC03108.1; -.			
DR	HSSB; 062226; IVH.			
DR	InterPro; IPR0011757; Hedgehog_hint.			
DR	InterPro; IPR003586; Hedgehog_hinct.			
DR	InterPro; IPR003587; Hedgehog_hinct.			
DR	InterPro; IPR003320; HH_signal.			
DR	InterPro; IPR002375; Pr/PY_rp_transf.			
DR	InterPro; IPR001657; SonicHH.			
DR	Pfam; PF01085; HH_signal; 1.			
DR	Pfam; PF01079; Hint; 1.			
DR	PRINTS; PR00632; SONICHHOG.			
DR	Prodom; PD003042; HH_signal; 1.			
DR	SMART; SM00305; Hintc; 1.			
DR	SMART; SM00306; Hintn; 1.			
DR	PROSITE; PS00103; PUR_PYR_TRANSFER; 1.			

Best Local Similarity 53.7%; Pred. No. 3e-82;
Matches 231; Conservative 53; Mismatches 119; Indels 27; Gaps 8

[illegible]

RESULT 6			
073803			
ID	073803	PRELIMINARY;	PRT; 442 AA.
AC	073803;		
DT	01-AUG-1998 (TEMBLrel. 07, Created)		
DT	01-AUG-1998 (TEMBLrel. 07, Last sequence update)		
DT	01-JUN-2002 (TEMBLrel. 21, Last annotation update)		
DE	Fugu hedgehog.		
GN	FHh.		
OS	Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neuteleostei		
OC	Acanthomorpha; Acanthopterygii; Perciformes; Tetraodoniformes;		
OC	Tetraodontidae; Takifugu.		
OX	NCBI_TaxID=31033;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=99177347; PubMed=10077531;		
RX	Gelliner K., Brenne S.;		
RT	"Analysis of 148 kb of genomic DNA around the wnt1 locus of Fugu		
RT	rubripes.";		
RL	Genome Res. 9:251-258(1999).		
DR	EMBL; AF056116; AAC34384.1; .		
DR	HSSP: Q62226; 1VHH.		
DR	InterPro: IPR001767; Hedgehog_hnt.		
DR	InterPro: IPR003586; Hedgehog_hntc.		
DR	InterPro: IPR003587; Hedgehog_hntv.		
DR	InterPro: IPR003207; HH_signal.		
DR	InterPro: IPR002203; Interf.		
DR	Pfam: PF01085; HH_signal; 1.		
DR	Pfam: PF01079; Hntc; 1.		
DR	ProDom: PD003042; HH_signal; 1.		
DR	SMART; SM00305; Hntc; 1.		
DR	SMART; SM00306; Hntv; 1.		
SO	SEQUENCE 442 AA; 49286 MW; 1CB2BA23BA8B748E8 CRC64;		
Query Match	46.5%;	Score 1063.5;	DB 13; Length 442;

Best Local Similarity 51.2%; Pred. No. 2.7e-77;
Matches 230; Conservative 60; Mismatches 128; Indels 31; Gaps 12;

[illegible]

RESULT 7			
Q9GRAB			
ID	Q9GRAB	PRELIMINARY;	PRF; 447 AA.
AC	Q9GRAB;		
DT	01-MAR-2001 (TREMBLrel. 16, Created)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Hedgehog protein (Fragment).		
OS	Gryllus bimaculatus (Two-spotted cricket).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae		
OC	Gryllinae; Gryllus.		
NCBI_TaxID=6999;			
[1]			
SEQUENCE FROM N.A.			
RA	MEDLINE-20461145; PubMed-11003837;		
RA	Niwa N., Inoue Y., Nozawa A., Saito M., Mismari Y., Ohuchi H.,		
RA	Yoshioke H., Noji S.;		
RT	"Correlation of diversity of leg morphology in Gryllus bimaculatus		
RT	(cricket) with divergence in dpp expression pattern during leg		
RT	development.";		
RL	Development 127:4373-4381(2000).		
EMBL	AB044709; BAB19658.1; -.		
HSSP	Q62226; 1VNH.		
DR	InterPro: IPR001767; Hedgehog_hnt.		
DR	InterPro: IPR003586; Hedgehog_hntC.		
DR	InterPro: IPR003587; Hedgehog_hntN.		
DR	InterPro: IPR000320; HH_signal.		
DR	InterPro: IPR001657; SonicHh.		
DR	Pfam: PF01085; HH_signal; 1.		
DR	Pfam: PF01079; Hnt; 1.		
DR	PRINTS: PRO0632; SONICHHOG.		
DR	ProDom: PD003042; HH_signal; 1.		
DR	SMART: SM00305; HntC; 1.		
DR	SMART: SM00306; HntN; 1.		
FT	NON_TER		
Q0	SEQUENCE 447 AA; 4808 MW; 29AEFB061C3EEF0 CRC64;		

Db 121 TWHLEGGTKLVKDLSPGDRVLAADQGRILYSDFLTFLDR 161

RESULT 10

OC 09XGV7 PRELIMINARY; PRT; 150 AA.
 AC 09XGV7;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Sonic hedgehog (Fragment).
 OS Ambystoma mexicanum (Axolotl).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandroidae; Ambystomatidae;
 OC Ambystoma.
 OC NCBI_TaxID=8296;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Torok M.A., Ispizua-Belmonte J.C., Gardiner D.M., Bryant S.V.;
 RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF031480; AAD18128.1; --
 DR HSSP; Q62226; 1VHH.
 DR InterPro; IPR001767; Hedgehog_hint.
 DR InterPro; IPR000320; HH_signal.
 DR InterPro; IPR001657; SonicHH.
 DR Pfam; PF01085; HH_signal; 1.
 DR Pfam; PF01079; Hnt; 1.
 DR PRINTS; PR00632; SONICHHOG.
 DR ProDom; PD003042; HH_signal; 1.
 FT NON_TER 1 150
 FT NON_TER 1 150
 SQ SEQUENCE 150 AA; 16599 MW; 986F65037A69A8AC CRC64;

Query Match 32.2%; Score 736; DB 13; Length 150;
 Best Local Similarity 93.3%; Pred. No. 1.3e-51;
 Matches 140; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OC 84 DIFFDEENTGADRLMTORCKDKLNALAISSVNMOPGVRRLRYTEGDEDEHHSEESLHAE 143
 DB 1 DIFFDEENTGADRLMTORCKDKLNALAISSVNMOPGVRRLRYTEGDEDEHHSEESLHAE 60
 QY 144 GRAVDITTSDBRDSKYGMLARLAVEGFPMVYVESKAHHCYSKAKNSVAAKSGGCPFS 203
 DB 61 GRAVDITTSDBRDSKYGMLARLAVEGFPMVYVESKAHHCYSKAKNSVAAKSGGCPFS 120
 QY 204 ATYHLEGGTKLVKDLSPGDRVLAADQGR 233
 DB 121 AKVTLEHGVTRPVKDLRPDRVLAADQGR 150

RESULT 11
 OC 09XSI6 PRELIMINARY; PRT; 139 AA.
 AC 09XSI6;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Sonic hedgehog (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tissue-Tooth Germ;
 RA Koyama E., Iwamoto M., Ohmori T., Kurisu K., Wu C., Ookura T.,
 RA Baehrl M.M., Tucker T., Pacifici M.;
 RT "Development of Stratum Intermedium and its Role as a Sonic Hedgehog-
 Signaling Structure During Odontogenesis.";
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
 DB EMBL; AF144100; AAD33926.1; --

DR HSSP; Q62226; 1VHH.
 DR InterPro; IPR000320; HH_signal.
 DR InterPro; IPR001657; SonicHH.
 DR Pfam; PF01085; HH_signal; 1.
 DR PRINTS; PR00632; SONICHHOG.
 DR ProDom; PD003042; HH_signal; 1.
 FT NON_TER 1 139
 FT NON_TER 1 139
 SQ SEQUENCE 139 AA; 15961 MW; C83B0B96982046B1 CRC64;

Query Match 32.0%; Score 732; DB 6; Length 139;
 Best Local Similarity 98.6%; Pred. No. 2.5e-51;
 Matches 137; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OC 44 AVKQFIPNVAEKTIGASGRYEKTRNSERFELTPVNPDIIFKDEENTGADRLMTORC 103
 DB 1 AVKQFIPNVAEKTIGASGRYEKTRNSERFELTPVNPDIIFKDEENTGADRLMTORC 60
 QY 104 KDKLNALAISSVNMOPGVRRLRYTEGDEDEHHSEESLHAEGRAVDITTSDBRDSKYGMLA 163
 DB 61 KDKLNALAISSVNMOPGVRRLRYTEGDEDEHHSEESLHAEGRAVDITTSDBRDSKYGMLA 120
 QY 164 RLAVEAGFDWVYVESKAH 182
 DB 121 RLAVEAGFDWVYVESKAH 139

RESULT 12

OC 09W6C1 PRELIMINARY; PRT; 138 AA.
 AC 09W6C1;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Sonic hedgehog protein (Fragment).
 OS SHH.
 OS Eleutherodactylus coqui (Puerto Rican coqui).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidae; Leptodactylidae;
 OC Eleutherodactylus.
 OC NCBI_TaxID=57060;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Carl T.F., Richardson M.K., Olsson L., Schlosser G., Klimkowski M.W.,
 RA Hanken J.;
 RT "Differences in vertebrate limb development revealed by studies of the
 RT direct developing frog E. coqui.";
 RL Submitted (Dec-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF113403; AAD23436.1; --
 DR HSSP; Q62226; 1VHH.
 DR InterPro; IPR000320; HH_signal.
 DR InterPro; IPR001657; SonicHH.
 DR Pfam; PF01085; HH_signal; 1.
 DR PRINTS; PR00632; SONICHHOG.
 DR ProDom; PD003042; HH_signal; 1.
 FT NON_TER 1 138
 FT NON_TER 1 138
 SQ SEQUENCE 138 AA; 15751 MW; FF4156A17F4681F0 CRC64;

Query Match 31.0%; Score 710; DB 13; Length 138;
 Best Local Similarity 96.4%; Pred. No. 1.4e-49;
 Matches 133; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OC 49 IPNVAEKTLGASGRYEKTRNSERFELTPVNPDIIFKDEENTGADRLMTORCKDKLN 108
 DB 1 IPNVAEKTLGASGRYEKTRNSERFELTPVNPDIIFKDEENTGADRLMTORCKDKLN 60
 QY 109 ALAISSVNMOPGVRRLRYTEGDEDEHHSEESLHAEGRAVDITTSDBRDSKYGMLARLAVE 168
 DB 61 ALAISSVNMOPGVRRLRYTEGDEDEHHSEESLHAEGRAVDITTSDBRDSKYGMLARLAVE 120
 QY 169 AGFDWVYVESKAHHCYSV 186
 DB 121 AGFDWVYVESKAHHCYSV 139

```

Db 121 AGEDWYTESKAHHCY 138

RESULT 13
09WV29 PRELIMINARY; PRT; 177 AA.
AC 09WV29;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Indian hedgehog protein (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RA Garages P.L., Meyer R.A., Jr., Brown C.A., Price D.K.;
RT "Indian hedgehog in rat.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF162914; AAD45372.1; -.
DR HSSP: Q62226; 1VHH.
DR MEROPS: C46.003; -.
DR InterPro: IPR001767; Hedgehog_hnt.
DR InterPro: IPR003587; Hedgehog_hntn.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR001657; SonICHH.
DR Pfam: PF01085; HH_signal; 1.
DR Pfam: PF01079; Hntc; 1.
DR PRINTS: PR00632; SONICHOG.
DR PRODOM: PD003042; HH_signal; 1.
DR SMART: SM00306; Hntn; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 177 AA; 19739 MW; CA5871626A565FE5 CRC64;

Query Match 30.5%; Score 698; DB 11; Length 177;
Best Local Similarity 73.4%; Pred. No. 1.9e-48;
Matches 130; Conservative 17; Mismatches 30; Indels 0; Gaps 0;

QY 98 LMTORCKDKLNALAIYVNMOPGVRRLVTEGWDDEGHSESLHYEGRAVDITTSDDR 157
DB 1 LMTORCKDKRLSLAIYVNMOPGVRRLVTEGWDDEGHSESLHYEGRAVDITTSDDR 60
QY 158 KYGMALARAVAGFDWYTESKAHHCYVKAENSVAAKSGCGFPGSATVHLEOGCTKL 217
DB 61 KYGLLARAVAGFDWYTESKAHHCYVKAENSVAAKSGCGFPGAGOVHLETGERVALS 120
QY 218 DLPRGDRVLADDOGRLLYSDFLFLDDEGAKKVFYVETLEPRERLLLTAAHL 274
DB 121 AVKPGDRVLAMGEDGNPFESDVLFLDDEPRRLRAFGVETDPPRRLLTAAHL 177

RESULT 14
096699 PRELIMINARY; PRT; 185 AA.
AC 096699;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Hedgehog protein (Fragment).
GN HH.
OS Junonia coenia (Peacock butterfly) (Precis coenia).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Plekoptera; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditylata;
OC Papilionidae; Nymphalidae; Nymphalinae; Junonia.
OX NCBI_TaxID=39708;
RN [1]
RP SEQUENCE FROM N.A.
RA Keys D.N., Lewis D.J., Selegue J.E., Pearson B.J., Goodrich L.V.,
RA Johnson R.L., Gates J., Scott M.P., Carroll S.B.;
RT "Recruitment of a hedgehog regulatory circuit in butterfly eyespot

```

```

RT evolution.";
RL Science 0:0-0(1999).
DR EMBL: AF117742; AAD08931.1; -.
DR HSSP: Q62226; 1VHH.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR001657; SonICHH.
DR Pfam: PF01085; HH_signal; 1.
DR PRINTS: PR00632; SONICHOG.
DR PRODOM: PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 185 AA; 20745 MW; E9B9F4CE2FAE662 CRC64;

Query Match 27.4%; Score 628; DB 13; Length 185;
Best Local Similarity 97.4%; Pred. No. 8.9e-43;
Matches 118; Conservative 20; Mismatches 30; Indels 2; Gaps 2;

```

```

QY 59 ASGRTEKTRNSERFELTPNPNPDIIFKDENTGADRLMTORCKDKLNALAIYVNM 118
DB 3 ASGPPEGRITRDEKFDLPVNPNDIDFDEGTGADRLMTORCKDKLNALAIYVNM 62
QY 119 PGVRLRVTEGWDDEGHSESLHYEGRAVDITTSDDRSKYGMALARAVAGFDWY 178
DB 63 PGVRLRVTEGWDDESHLDSLHYEGRAVDITTSDDRSHKGMALARAVAGFDWY 122
QY 179 KAHHCYVKAENSVAAKSGCGFPGSATVHLEOGCTKLVRKDLPRGDRVLAA 228
DB 123 RSYHCYVKTSSVGT-GAGCFPGAVVHTNGPXD-IASLKGKAYLAA 170

RESULT 15
080GN4 PRELIMINARY; PRT; 99 AA.
AC 080GN4;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Sonic hedgehog (Fragment).
OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphi; Galeidae; Carcharhiniformes;
OC Scyllorhinidae; Scyllorhinus.
OX NCBI_TaxID=7830;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21930458; Pubmed-11932743;
RA Tanaka M., Munsterberg A., Anderson W.G., Prescott A.R., Hazon N.,
RA Tickle C.;
RT "Fin development in a cartilaginous fish and the origin of vertebrate
limbs.";
RL Nature 416:527-531(2002).
DR EMBL: AF393835; AAM08228.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 99 AA; 11468 MW; 831864BDB34A1A20 CRC64;

Query Match 23.1%; Score 528; DB 13; Length 99;
Best Local Similarity 97.0%; Pred. No. 4e-35;
Matches 96; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 79 PNYPNDIIFKDENTGADRLMTORCKDKLNALAIYVNMOPGVRRLVTEGWDDEGH 138
DB 1 PNYPNDIIFKDENTGADRLMTORCKDKLNALAIYVNMOPGVRRLVTEGWDDEGH 60
QY 139 SLHYEGRAVDITTSDDRSKYGMALARAVAGFDWY 177
DB 61 SLHYEGRAVDITTSDDRRTKGMALARAVAGFDWY 99

```

Search completed: February 20, 2003, 10:11:51
 Job time : 29.1092 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:05:32 ; Search time 34.4615 Seconds

(without alignments)
1836.661 Million cell updates/sec

Title: us-09-827-110a-15

Perfect score: 2469 1 MLTARCLLVVSSILVCS.....GNAVKSSXSRGAGGAREGA 475

Sequence: BLOSUM62

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : A_Geneseq_101002.*

```
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2467	99.9	475	16	AA77341 Human sonic hedgeh
2	2467	99.9	475	20	AA70585 Human sonic hedgeh
3	2467	99.9	475	20	AA70515 Human sonic hedgeh
4	2467	99.9	475	20	AA77770 Human sonic hedgeh
5	2467	99.9	475	21	AA794473 Human Shh hedgehog
6	2467	99.9	475	21	AA795977 Human sonic hedgehog
7	2467	99.9	475	21	AA795286 Human sonic hedgehog
8	2467	99.9	475	21	AA796248 Partial human Shh.
9	2467	99.9	475	21	AA770681 Human sonic hedgehog
10	2467	99.9	475	22	AA65748 Human sonic hedgehog

11	2467	99.9	475	22	AA84674 Amino acid sequenc
12	2467	99.9	475	22	AA804687 Human sonic hedgeh
13	2467	99.9	475	22	AA805377 Human sonic hedgeh
14	2467	99.9	475	22	AA831222 Amino acid sequenc
15	2467	99.9	475	22	AA860265 Human Sonic hedgeh
16	2467	99.9	475	23	AA799483 Human sonic hedgeh
17	2467	99.9	475	23	AA879137 Human sonic hedgeh
18	2467	99.9	475	23	AA820923 Human Shh protein
19	2467	99.9	475	23	AA814297 Human sonic hedgeh
20	2467	99.9	475	23	AA85085 Human sonic hedgeh
21	2407	97.5	462	22	AA85738 Human sonic hedgeh
22	2403	97.3	462	19	AA848736 Human mutated sonl
23	2401	97.2	462	19	AA848735 Human mutated sonl
24	2068	83.8	437	19	AA861488 Mouse sonic hedgeh
25	2068	83.8	437	22	AA809292 Mouse sonic hedgeh
26	2068	83.8	437	22	AA85086 Mouse sonic hedgeh
27	2068	83.8	437	22	AA797559 Mouse sonic hedgeh
28	2068	83.8	437	23	AA821264 Mouse sonic hedgeh
29	2065	83.6	437	16	AA87339 Mouse sonic hedgeh
30	2065	83.6	437	20	AA70587 Mouse sonic hedgeh
31	2065	83.6	437	20	AA70513 Mouse sonic hedgeh
32	2065	83.6	437	20	AA79768 Mouse sonic hedgeh
33	2065	83.6	437	20	AA794471 Mouse Shh hedgehog
34	2065	83.6	437	21	AA795975 Mouse sonic hedgeh
35	2065	83.6	437	21	AA795284 Mouse sonic hedgeh
36	2065	83.6	437	21	AA796246 Partial mouse Shh.
37	2065	83.6	437	21	AA70679 Mouse Sonic hedgeh
38	2065	83.6	437	22	AA65746 Mouse sonic hedgeh
39	2065	83.6	437	22	AA85736 Mouse sonic hedgeh
40	2065	83.6	437	22	AA84672 Amino acid sequenc
41	2065	83.6	437	22	AA804687 Mouse sonic hedgeh
42	2065	83.6	437	22	AA805375 Mouse sonic hedgeh
43	2065	83.6	437	22	AA831220 Amino acid sequenc
44	2065	83.6	437	22	AA860263 Mouse Sonic hedgeh
45	2065	83.6	437	23	AA799481 Mouse sonic hedgeh

ALIGNMENTS

RESULT 1	AA77341	
ID	AA77341 standard; Protein; 475 AA.	
XX	AA77341;	
AC	14-MAR-1996 (first entry)	
XX		
DT	Human sonic hedgehog protein.	
XX		
DE	Human sonic hedgehog protein; probe; primer; diagnostic;	
XX		
XX	Human sonic hedgehog protein; probe; primer; diagnostic;	
KW	nervous system disorder; gene therapy; antibody.	
KW		
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	Peptide	24..29
FT	Misc-difference	/note="conserved sequence (AA77341)"
FT		/note="unspecified amino acid"
XX		
PN	W09518856-A1.	
XX		
PD	13-JUL-1995.	
XX		
PF	30-DEC-1994;	94WO-US14992.
XX		
PR	14-DEC-1994;	94US-0356060.
XX		
PR	30-DEC-1993;	93US-0176427.
XX		
PA	(HARD) HARVARD COLLEGE.	
XX	(TWCR) IMPERIAL CANCER RES TECHNOLOGY.	

PI Ingham PW, McMahon AP, Tablin CJ;
 XX WPI, 1995-255060/33.
 DR N-PSDB; AA091639.
 XX
 PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
 PT to treat degenerative nervous system disorder(s) and in gene
 PT therapy.
 PS
 XX
 PS
 XX
 CC Claim 17; Page 143-45; 210pp; English.
 CC
 CC The sequence represents a human sonic hedgehog protein, homologous
 CC to a Drosophila hedgehog protein (AA073357), and is encoded by a cDNA
 CC isolated from a human fetal lung cDNA library. Probes and primers
 CC derived from the sonic hedgehog gene may be used as diagnostic
 CC agents for neuromuscular, autonomic or central nervous system
 CC disorders, and the gene may also be used in gene therapy.
 CC Antibodies generated from the protein may be used as therapeutic or
 CC research reagents.
 CC
 SQ Sequence 475 AA;

Query Match 99.9%; Score 2467; DB 16; Length 475;
 Best Local Similarity 100.0%; Pred. No. 5.9e-218;
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLLARCLLVVSSLLVCSGLACGPGRGFRHHPKLTPLAYKOFIPNVAEKLGLASG 60
 DB 1 MLLARCLLVVSSLLVCSGLACGPGRGFRHHPKLTPLAYKOFIPNVAEKLGLASG 60
 OY 61 RREGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNOMP 120
 DB 61 RREGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNOMP 120
 OY 121 KLRVTEGMEDGHSEESIHYEGRAVDITTSDRNSKYGLARLVAEAGFDWYTESKAH 180
 DB 121 KLRVTEGMEDGHSEESIHYEGRAVDITTSDRNSKYGLARLVAEAGFDWYTESKAH 180
 OY 181 HCSYVAENSVAKSGCGPGSATVHLEGGGTRKLVKDLSPGDRVLAADOGRLYSDFLT 240
 DB 181 HCSYVAENSVAKSGCGPGSATVHLEGGGTRKLVKDLSPGDRVLAADOGRLYSDFLT 240
 OY 241 FLDRODGAKKVYVITRPRERLLTAHLLFVAPHNDSATGEPASSGSGPPSGALG 300
 DB 241 FLDRODGAKKVYVITRPRERLLTAHLLFVAPHNDSATGEPASSGSGPPSGALG 300
 OY 301 PALFASRRVPGORVYVAERDGRRLPAVAVSYTLSEEAAGAVPLTAOGTILINRYL 360
 DB 301 PALFASRRVPGORVYVAERDGRRLPAVAVSYTLSEEAAGAVPLTAOGTILINRYL 360
 OY 361 ASCYAVIEEHSMAHRAFAFRLAHLLALAPARTDRGDSGGGGRGGRVALTAPGA 420
 DB 361 ASCYAVIEEHSMAHRAFAFRLAHLLALAPARTDRGDSGGGGRGGRVALTAPGA 420
 OY 421 AARPAGAGAGTGHWSQLLYOIGTWLIDSEALHPGLGMAYKSSXSRAGGAREGA 475
 DB 421 AARPAGAGAGTGHWSQLLYOIGTWLIDSEALHPGLGMAYKSSXSRAGGAREGA 475

RESULT 2
 AA05859
 ID AA05859 standard; Protein; 475 AA.
 AC AA05859;
 XX
 XX
 DT 02-AUG-1999 (first entry)
 DE Human Sonic hedgehog Shh protein.
 XX
 XX Sonic hedgehog; Shh; human; epithelial tissue; epithelium;
 KM cutaneous tissue; skin; hair; wound healing; vulvectomy;
 KM burn; skin grafting; pressure sore; ulcer; ulcerative colitis;
 KM alopecia; psoriasis; keratosis; acne; comedogenic lesion;

KW folliculitis; pseudofolliculitis; keratocanthoma; callosities;
 KM Darier's disease; scar; autoimmune disease; pemphigus;
 KM epidermolysis; lupus lesion; desquamative lesion; carcinoma;
 XX therapy; hedgehog therapeutic; ptc therapeutic; patched.
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 463
 FT /note= "encoded by NNN"

XX W09920298-A1.
 XX 29-APR-1999.
 XX
 XX 20-OCT-1998; 98MO-US22227.
 XX 11-SEP-1998; 98US-0151999.
 XX 20-OCT-1997; 97US-0955552.
 XX
 XX (ONTO-) ONTOGENY INC.
 XX
 XX
 PI Wang EA;
 XX
 DR WPI, 1999-288170/24.
 DR N-PSDB; AA025622.

Use of hedgehog polypeptides on patched therapeutics

Claim 26; Page 127-128; 146pp; English.

The present sequence represents human Sonic hedgehog protein Shh. The invention relates to a method for modulating the growth state an epithelial cell by ectopically contacting the epithelial cell, in vitro or in vivo, with a hedgehog therapeutic (i.e. a hedgehog polypeptide or gene therapy construct) or ptc therapeutic (i.e. a small organic molecule that mimics the effect of hedgehog proteins on patched signalling, or activates or potentiates patched signalling) in an amount effective to alter the rate of proliferation of the epithelial cell. The hedgehog therapeutic preferably comprises at least a bioactive extracellular portion of a hedgehog protein (see AA05854-62) encoded by a vertebrate hedgehog gene (see AA05617-25), especially a human hedgehog gene. Promotion of proliferation of epithelial cells can be used to control a wound healing process in e.g. burn treatment, skin regeneration, skin grafting, pressure sore treatment, dermal ulcer treatment, post surgery scar reduction or treatment of ulcerative colitis (claimed). It can also be used to induce hair growth for the treatment of alopecia (claimed). Inhibition of the growth of epithelial tissue can be used to treat or prevent hyperplastic or neoplastic conditions, e.g. psoriasis, keratosis, acne, comedogenic lesions, folliculitis and pseudofolliculitis, keratocanthoma, callosities, Darier's disease, keloids, hypertrophic scars, or autoimmune disorders, e.g. aphthous ulcers, pemphigus vulgaris, pemphigus foliaceus, pemphigus vegetans, pemphigus erythematous, epidermolysis, lupus lesions, desquamative lesions or carcinomas. The methods can also be used to counteract the effects of ageing on skin.

SQ Sequence 475 AA;
 Query Match 99.9%; Score 2467; DB 20; Length 475;
 Best Local Similarity 100.0%; Pred. No. 5.9e-218;
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLLARCLLVVSSLLVCSGLACGPGRGFRHHPKLTPLAYKOFIPNVAEKLGLASG 60
 DB 1 MLLARCLLVVSSLLVCSGLACGPGRGFRHHPKLTPLAYKOFIPNVAEKLGLASG 60
 OY 61 RREGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNOMP 120
 DB 61 RREGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNOMP 120

QY 121 KLRVTEGMDDEGHSESLHTEGRAVDITTSDRDSKYGMLARLAVAGEDWVYESKAH 180
 DB 121 KLRVTEGMDDEGHSESLHTEGRAVDITTSDRDSKYGMLARLAVAGEDWVYESKAH 180
 QY 181 IHCYKAKENSVAAKSGCGPESATVHLEOGGSKLVKDLSPEDRYLADDDGRLLYSPFLT 240
 DB 181 IHCYKAKENSVAAKSGCGPESATVHLEOGGSKLVKDLSPEDRYLADDDGRLLYSPFLT 240
 QY 241 FLDDDDGAKKVFYIETREPERELLTAAHLLEFAPHNDSATGPEASSGSGPPSGALG 300
 DB 241 FLDDDDGAKKVFYIETREPERELLTAAHLLEFAPHNDSATGPEASSGSGPPSGALG 300
 QY 301 PRALFASRVRRGQRYVVAERDGRLLPAAVHSVTLSSEAAAGAYAPLTAGTILLINRVL 360
 DB 301 PRALFASRVRRGQRYVVAERDGRLLPAAVHSVTLSSEAAAGAYAPLTAGTILLINRVL 360
 QY 361 ASCYAVIEEHSMAHRAFPRLAHLAALAPARTDRGDSGGSGDGGGGRVALTPAGA 420
 DB 361 ASCYAVIEEHSMAHRAFPRLAHLAALAPARTDRGDSGGSGDGGGGRVALTPAGA 420
 QY 421 ADAPGAGATAGIHWYSOLLVIGTWLDESLAHLPLGMAVKSXSXSGAGGARREGA 475
 DB 421 ADAPGAGATAGIHWYSOLLVIGTWLDESLAHLPLGMAVKSXSXSGAGGARREGA 475

RESULT 3 AAV05515

AAV05515 standard; Protein: 475 AA.

AC AAV05515:

DT 05-JUL-1999 (first entry)

DE Human Sonic hedgehog protein Shh.

XX Sonic hedgehog; Shh protein; human; hedgehog therapeutic;

KW ptc therapeutic; patched; signal transduction; muscle atrophy;

KM cohesin; muscular myopathy; myoblastic sarcoma; therapy.

XX Homo sapiens.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 463 /note= "encoded by NNN"

XX MO9910004-A2.

XX 04-MAR-1999.

XX 28-AUG-1998; 98WO-US17922.

XX 29-AUG-1997; 97US-0057394.

XX (ONTO-) ONTOGENY INC.

XX Bladgen CS, Currie PD, Hughes SM, Ingham PW;

XX WPI: 1999-243557/20.

XX N-PSDB: AAX25103.

XX A new method to regulate muscle growth

XX Disclosure; Page 118-120; 130pp; English.

XX The present sequence is human Sonic hedgehog protein Shh. The

CC invention relates to a method for modulating the formation and/or

CC maintenance of muscle tissue by ectopically contacting muscle

CC cells, especially muscle stem/progenitor cells, in vitro or in

CC vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and

CC gene therapy constructs) or ptc therapeutic (i.e. a small organic

CC molecule that mimics the effect of hedgehog proteins on patched

CC signalling, or activates or potentiates patched signalling) in an

CC amount effective to alter the growth state of the treated cells.

CC Also claimed is a method for treatment or prevention of disorders
 CC of, or surgical or cosmetic repair of, such muscle tissues, by
 CC administering a hedgehog polypeptide or ptc therapeutic. The
 CC disorder may be muscle atrophy, in particular skeletal muscle
 CC atrophy or cardiac muscle atrophy, cachexia, or muscular myopathy
 CC (all claimed). The hedgehog polypeptide or ptc therapeutic can
 CC inhibit growth of myoblastic-derived tissue to provide treatment of
 CC hyperplastic or neoplastic growth of muscle tissue such as in
 CC myoblastic sarcoma (also claimed). The hedgehog therapeutic
 CC preferably comprises at least a bioactive extracellular portion of
 CC a hedgehog protein (see AAV05510-19) encoded by a vertebrate hedgehog
 CC gene (see AAX25098-107), especially a human hedgehog gene.

CC Sequence 475 AA:

Query Match 99.9%; Score 2467; DB 20; Length 475;

Best Local Similarity 100.0%; Pred. No. 5.9e-218; Mismatches 0; Gaps 0;

Matches 475; Conservative 0; Indels 0; Gaps 0;

QY 1 MLTARCLLVLYSSLLVCSGLACPGRGFGKRRHPKRLPLAYKQPIPNVAEKTLAGSG 60

DB 1 MLTARCLLVLYSSLLVCSGLACPGRGFGKRRHPKRLPLAYKQPIPNVAEKTLAGSG 60

QY 61 RYEGKISNSERFKEKLPNTNPDIIFKDEENTGADRLMTQCKDKNALAISVNMOPGV 120

DB 61 RYEGKISNSERFKEKLPNTNPDIIFKDEENTGADRLMTQCKDKNALAISVNMOPGV 120

QY 121 KLRVTEGMDDEGHSESLHTEGRAVDITTSDRDSKYGMLARLAVAGEDWVYESKAH 180

DB 121 KLRVTEGMDDEGHSESLHTEGRAVDITTSDRDSKYGMLARLAVAGEDWVYESKAH 180

QY 181 IHCYKAKENSVAAKSGCGPESATVHLEOGGSKLVKDLSPEDRYLADDDGRLLYSPFLT 240

DB 181 IHCYKAKENSVAAKSGCGPESATVHLEOGGSKLVKDLSPEDRYLADDDGRLLYSPFLT 240

QY 241 FLDDDDGAKKVFYIETREPERELLTAAHLLEFAPHNDSATGPEASSGSGPPSGALG 300

DB 241 FLDDDDGAKKVFYIETREPERELLTAAHLLEFAPHNDSATGPEASSGSGPPSGALG 300

QY 301 PRALFASRVRRGQRYVVAERDGRLLPAAVHSVTLSSEAAAGAYAPLTAGTILLINRVL 360

DB 301 PRALFASRVRRGQRYVVAERDGRLLPAAVHSVTLSSEAAAGAYAPLTAGTILLINRVL 360

QY 361 ASCYAVIEEHSMAHRAFPRLAHLAALAPARTDRGDSGGSGDGGGGRVALTPAGA 420

DB 361 ASCYAVIEEHSMAHRAFPRLAHLAALAPARTDRGDSGGSGDGGGGRVALTPAGA 420

QY 421 ADAPGAGATAGIHWYSOLLVIGTWLDESLAHLPLGMAVKSXSXSGAGGARREGA 475

DB 421 ADAPGAGATAGIHWYSOLLVIGTWLDESLAHLPLGMAVKSXSXSGAGGARREGA 475

RESULT 4
AAW97770

ID AAW97770 standard; Protein: 475 AA.

AC AAW97770:

DT 21-MAY-1999 (first entry)

DE Human Sonic hedgehog (Shh) protein.

XX Sonic hedgehog; Shh protein; human; dopaminergic; GABA-ergic;

KW ptc therapeutic; patched; signal transduction; Parkinson's disease;

KM Huntington's disease; amyotrophic lateral sclerosis;

OS cerebral ischemia; hypoxia; neuroprotective; therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 463 /note= "encoded by NNN"

XX A new method to regulate muscle growth

XX Disclosure; Page 118-120; 130pp; English.

XX The present sequence is human Sonic hedgehog protein Shh. The

CC invention relates to a method for modulating the formation and/or

CC maintenance of muscle tissue by ectopically contacting muscle

CC cells, especially muscle stem/progenitor cells, in vitro or in

CC vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and

CC gene therapy constructs) or ptc therapeutic (i.e. a small organic

CC molecule that mimics the effect of hedgehog proteins on patched

CC signalling, or activates or potentiates patched signalling) in an

CC amount effective to alter the growth state of the treated cells.

PN W09904775-A2.
 XX 04-FEB-1999.
 XX 24-JUL-1998; 98WO-US15419.
 XX 24-JUL-1997; 97US-0900220.
 XX (ONTO-) ONTOGENY INC.
 XX Mahanthappa NK, Miao N, Pang K, Wang M;
 XX WPI, 1999-142578/12.
 DR N-PSDB; AAX07276.
 XX
 PT Increasing the survival of neuronal, dopaminergic and GABA-nergic
 PT cells - by using a ptc therapeutic such as a protein kinase
 PT inhibitor, or an agent derived from hedgehog polypeptides, useful in
 PT the treatment of Parkinson's disease
 XX
 PS Disclosure; Page 93-95; 138pp; English.
 XX
 CC This polypeptide is human Shh Sonic hedgehog protein. The
 CC invention is based on the finding that hedgehog proteins are useful
 CC as protective agents in the treatment and prophylaxis of
 CC neurodegenerative disorders resulting from the loss of dopaminergic
 CC and/or GABA-nergic neurons, or the general loss of tissue from the
 CC substantia nigra. Exemplary disorders include Parkinson's disease,
 CC Huntington's disease (both claimed), amyotrophic lateral sclerosis
 CC and cerebral ischaemia. The invention relates to hedgehog
 CC therapeutics (i.e. hedgehog polypeptides and gene therapy constructs
 CC e.g. constructs encoding recombinant hedgehog polypeptides and
 CC trans-activation constructs for altering hedgehog gene regulatory
 CC sequences) and ptc therapeutics (i.e. agents which mimic the effect
 CC of naturally occurring hedgehog proteins on patched signalling)
 CC that are effective in both human and animal subjects. Human Ihh
 CC and Dhh polypeptides (see AAM97763-64) are preferred. The products
 CC can also be used for the maintenance of differentiated neurons in
 CC cultures, and to enhance the implantation of such neuronal cells in
 CC an animal. They can be used to prevent or treat neurodegenerative
 CC conditions arising from the use of certain drugs, and in the
 CC prevention and/or treatment of hypoxia, e.g. as a neuroprotective
 CC agent.
 XX
 XX Sequence 475 AA:
 SQ
 Query Match 99.9%; Score 2467; DB 20; Length 475;
 Best Local Similarity 100.0%; Pred. No. 5.9e-218;
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 361 ASCYAVIEESHMAHRAFAEFLAHLAALAPARTDRGDSGGGGRVATAPGA 420
 DB 361 ASCYAVIEESHMAHRAFAEFLAHLAALAPARTDRGDSGGGGRVATAPGA 420
 OY 421 ADAPGAGATAGIHWSQOLLYOIGTWLDBSEALHPLGMAVKSXSRGAGGAREGA 475
 DB 421 ADAPGAGATAGIHWSQOLLYOIGTWLDBSEALHPLGMAVKSXSRGAGGAREGA 475
 RESULT 5
 AAM94473
 ID AAM94473 standard; Protein; 475 AA.
 XX
 AC AAM94473;
 XX
 DT 29-APR-1999 (first entry)
 XX
 DE Human Shh hedgehog protein sequence.
 XX
 KW patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;
 KW brain infarction; cerebral infarction; transient ischaemic attack;
 KW stroke; cerebral infarct volume; spinal cord; oedema; trauma;
 KW haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.
 XX
 OS Homo sapiens.
 XX
 FT Key location/Qualifiers
 FT Misc-difference 463 /label= unknown
 FT /note= "encoded by NNN"
 XX
 W09900117-A2.
 XX
 07-JAN-1999.
 XX
 26-JUN-1998; 98WO-US13387.
 XX
 27-JUN-1997; 97US-0883656.
 XX
 PA (ONTO-) ONTOGENY INC.
 XX
 XX Mahanthappa NK;
 PT
 DR WPI; 1999-095458/08.
 DR N-PSDB; AAX16187.
 XX
 PT Method for limiting damage to neurons caused by ischaemic or epoxic
 PT conditions - is used for the treatment and prevention of e.g.
 PT cerebral infarction, stroke and transient ischaemic attacks
 XX
 PS Disclosure; Page 72-74; 104pp; English.
 XX
 CC A method has been developed for limiting the damage to neuronal cells by
 CC ischaemic or epoxic conditions by administering a ptc (patched)
 CC therapeutic agent to reduce cerebral infarct volume (CIV). Damage to
 CC neuronal cells can also be limited by administering a gene activation
 CC construct which recombines with the genomic hedgehog gene to provide a
 CC heterologous transcription regulator linked to the coding region of this
 CC gene. Administration of the ptc therapeutic agent is used to protect
 CC cerebral tissues against ischaemic injury; to treat cerebral infarct or
 CC ischaemia, stroke (thrombotic or embolic) and transient ischaemic
 CC attacks. It may also be used as a prophylactic in many other cases of
 CC injury to the brain or spinal cord, oedema caused by trauma, haemorrhage
 CC and encephalomyelitis, or in conjunction with (coronary bypass) surgery.
 CC Treatment (which may be prophylactic) is used where ischaemic/epoxic
 CC conditions may cause cerebral hypoxia, or progressive loss of neurons
 CC due to oxygen depletion, including in patients with hypotension. The
 CC treatment reduces CIV by at least 25, particularly at least 70%. The
 CC present sequence represents a hedgehog sequence given in the present
 CC invention.
 XX
 XX Sequence 475 AA:
 SQ

Query Match 99.9%; Score 2467; DB 20; Length 475;
 Best Local Similarity 100.0%; Pred. No. 5.9e-218;
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVLSLVCSGLACPGRCGRRHPPKLTPLAYKQFIPNVAEKTGLASG 60
 D 1 MLLARCLLLVLSLVCSGLACPGRCGRRHPPKLTPLAYKQFIPNVAEKTGLASG 60
 QY 61 RYEGKISNSERFELPNYPNDIIFKDEENTGADRLMTORCKKLNALAI SVMNQPGV 120
 D 61 RYEGKISNSERFELPNYPNDIIFKDEENTGADRLMTORCKKLNALAI SVMNQPGV 120
 QY 121 KLRVTEGDEGHHSESLHYEGRAVDITTSDRRSKYGMRLARLAVEGPDWYYESKAH 180
 D 121 KLRVTEGDEGHHSESLHYEGRAVDITTSDRRSKYGMRLARLAVEGPDWYYESKAH 180
 QY 181 IHCSVKAENSVAAKSGCGFPGSATVHLEOGCTKLVDLSPDRVLAADDGRLLYSDFLT 240
 D 181 IHCSVKAENSVAAKSGCGFPGSATVHLEOGCTKLVDLSPDRVLAADDGRLLYSDFLT 240
 QY 241 FLDDDDGAKKVFYIETREPERELLTAHLLFYAPHNDSATGEPBASSGSPSGCALG 300
 D 241 FLDDDDGAKKVFYIETREPERELLTAHLLFYAPHNDSATGEPBASSGSPSGCALG 300
 QY 301 PRALFASVRPGORVYVAERDGRRLTPAAVHSVTLSEEAAGAYAPLTAGTLLINRVL 360
 D 301 PRALFASVRPGORVYVAERDGRRLTPAAVHSVTLSEEAAGAYAPLTAGTLLINRVL 360
 QY 361 ASCAIVAEHSWHAHAFAPRLAHLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420
 D 361 ASCAIVAEHSWHAHAFAPRLAHLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420
 QY 421 ADAGAGATAGIHMYSOGLIOTWLDSEALHPLGMAVKSXSRGAGGARBEA 475
 D 421 ADAGAGATAGIHMYSOGLIOTWLDSEALHPLGMAVKSXSRGAGGARBEA 475

RESULT 6
 AAY95977
 ID AAY95977 standard; Protein: 475 AA.
 AC AAY95977;
 DT 05-DEC-2000 (first entry)
 DE Human Sonic hedgehog Shh protein.
 XX
 KW Sonic hedgehog; Shh; human; agonist; antagonist;
 KW lipid modulator; vacuole; cholesterol disorder; lipid disorder;
 KW lipid metabolism; lipid storage; lipid transport; apolipoprotein;
 KW triglyceride; hypercholesterolemia; abetalipoproteinemia;
 KW hypobetalipoproteinemia; chylomicron retention; Anderson's disease;
 KW fat absorption; atherosclerosis; obesity; weight loss;
 KW vitamin A disorder; vitamin E disorder; anillopenia; anorectic;
 KW antiarteriosclerotic; gene therapy; diagnosis.
 KW
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 463 /note= "encoded by NNN"
 XX
 PN MO200051628-A2.
 PD 08-SEP-2000.
 XX
 PF 03-MAR-2000; 2000WO-US05662.
 XX
 PR 03-MAR-1999; 99US-0122640.
 PR 15-MAR-1999; 99US-0124446.
 XX
 PA (BIOU) BIOGEN INC.

XX Burkly L, Wang LC;
 PI WPI: 2000-611340/58.
 XX N-PSDB; AAA95977.
 DR
 XX
 PT use of lipid modulators (e.g. hedgehog agonists or antagonists) for
 PT modulating lipid metabolism and storage, especially useful for treating
 PT lipid metabolism or cholesterol disorders, e.g. obesity or
 PT hypercholesterolemia -
 PS Disclosure; Page 118-120; 136pp; English.
 XX
 CC The present sequence of that of human Sonic hedgehog (Shh)
 CC protein. The invention provides claimed methods for modulating
 CC lipid metabolism, for modulating vacuole formation in intestinal
 CC epithelial cells, for modulating the accumulation of fat in
 CC intestinal epithelial cells, for treating a cholesterol disorder
 CC and for treating a lipid metabolism disorder in an animal
 CC (especially a human) by administering a lipid modulator selected
 CC from a hedgehog antagonist or hedgehog agonist. In particular, the
 CC lipid metabolism disorder is a lipid storage disorder, a lipid
 CC transport disorder, an apolipoprotein disorder, a triglyceride
 CC disorder, e.g. a triglyceride metabolism disorder, a triglyceride
 CC transport disorder or a triglyceride storage disorder; a
 CC diet-induced hypercholesterolemia, hypercholesterolemia,
 CC abetalipoproteinemia, hypobetalipoproteinemia; a
 CC chylomicron-retention disorder, Anderson's disease, a fat
 CC absorption disorder, e.g. obesity or associated with weight loss,
 CC normotriglyceridemic abetalipoproteinemia, an apolipoprotein-B100
 CC deficiency, a fat soluble vitamin disorder, where the fat soluble
 CC vitamin is vitamin A or E, or atherosclerosis (all claimed). The
 CC hedgehog antagonist binds to the hedgehog receptor, but does not
 CC elicit a response. It is preferably a hedgehog mimetic, a modified
 CC hedgehog protein, e.g. an inactive hedgehog variant, or an
 CC anti-hedgehog homologue, especially a human, chimeric or humanised
 CC antibody. The methods are useful in preventing these disorders or
 CC protecting a subject from these disorders. The hedgehog antagonist
 CC and agonist are also useful in diagnosis and research associated
 CC with these disorders. The lipid modulators may also be used as a
 CC part of a gene therapy protocol to deliver polynucleotides encoding
 CC these lipid modulators.
 CC
 CC Sequence 475 AA;
 XX
 SQ

Query Match 99.9%; Score 2467; DB 21; Length 475;
 Best Local Similarity 100.0%; Pred. No. 5.9e-218;
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVLSLVCSGLACPGRCGRRHPPKLTPLAYKQFIPNVAEKTGLASG 60
 D 1 MLLARCLLLVLSLVCSGLACPGRCGRRHPPKLTPLAYKQFIPNVAEKTGLASG 60
 QY 61 RYEGKISNSERFELPNYPNDIIFKDEENTGADRLMTORCKKLNALAI SVMNQPGV 120
 D 61 RYEGKISNSERFELPNYPNDIIFKDEENTGADRLMTORCKKLNALAI SVMNQPGV 120
 QY 121 KLRVTEGDEGHHSESLHYEGRAVDITTSDRRSKYGMRLARLAVEGPDWYYESKAH 180
 D 121 KLRVTEGDEGHHSESLHYEGRAVDITTSDRRSKYGMRLARLAVEGPDWYYESKAH 180
 QY 181 IHCSVKAENSVAAKSGCGFPGSATVHLEOGCTKLVDLSPDRVLAADDGRLLYSDFLT 240
 D 181 IHCSVKAENSVAAKSGCGFPGSATVHLEOGCTKLVDLSPDRVLAADDGRLLYSDFLT 240
 QY 241 FLDDDDGAKKVFYIETREPERELLTAHLLFYAPHNDSATGEPBASSGSPSGCALG 300
 D 241 FLDDDDGAKKVFYIETREPERELLTAHLLFYAPHNDSATGEPBASSGSPSGCALG 300
 QY 301 PRALFASVRPGORVYVAERDGRRLTPAAVHSVTLSEEAAGAYAPLTAGTLLINRVL 360
 D 301 PRALFASVRPGORVYVAERDGRRLTPAAVHSVTLSEEAAGAYAPLTAGTLLINRVL 360

QY 361 ASCYAVIEEHSNAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGA 420
 DB 361 ASCYAVIEEHSNAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGA 420
 QY 421 ADAPGAGATAGIHWYSOLLYOIGTWLIDSEALHPLGMAYKSSXSRCAGGAREGA 475
 DB 421 ADAPGAGATAGIHWYSOLLYOIGTWLIDSEALHPLGMAYKSSXSRCAGGAREGA 475

RESULT 7
 AAY95286
 ID AAY95286 standard; Protein; 475 AA.
 AC AAY95286;
 DT 12-SEP-2000 (first entry)
 DE Human Sonic hedgehog Shh protein.
 KW Sonic hedgehog; Shh; human; excitotoxicity; Parkinson's disease;
 KW Huntington's disease; neuronal degeneration; neuroprotective;
 KW dopaminergic; GABAergic; substantia nigra; therapy.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc-difference 463 /note= "encoded by NNN"
 FT WO200035948-A1.
 PN 22-JUN-2000.
 PD 03-DEC-1999; 99WO-US28721.
 PE 03-DEC-1999; 99WO-US28721.
 PR 03-DEC-1999; 98WO-US25676.
 PR 27-JAN-1999; 99US-0238243.
 PR 03-JUN-1999; 99US-0325602.
 PA (BIOJ) BIOGEN INC.
 PA (ONTO-) ONTOGENY INC.
 PI Galdes A, Mahanthappa N;
 DR WPI: 2000-431570/37.
 DR N-PSDB: AAA27881.
 PT Treating disorders involving excitotoxicity, e.g. trauma, hypoglycemia,
 PT senile dementia and Korsakoff's disease, by using lipophilic modified
 PT hedgehog polypeptide -
 PS Disclosure; Page 157-159; 174pp; English.
 XX The present sequence of that of human Sonic hedgehog (Shh)
 CC protein. The invention relates to a method for promoting the
 CC survival and/or functional performance of neuronal cells,
 CC especially substantia nigra, dopaminergic or GABAergic neurons that
 CC are susceptible to excitotoxicity, by contacting the cells, in vitro
 CC or in vivo, with a lipophilic (e.g. cholesterol) modified hedgehog
 CC polypeptide. The method is used to treat or prevent Parkinson's
 CC disease, Huntington's disease, domoic acid poisoning, spinal cord
 CC trauma, hypoglycemia, mechanical trauma to the nervous system,
 CC senile dementia, Korsakoff's disease, schizophrenia, AIDS dementia,
 CC multi-infarct dementia, mood disorders, depression, chemical
 CC toxicity, neuronal damage associated with uncontrolled seizures
 CC such as epileptic seizures, neuronal injury associated with HIV and
 CC AIDS, neurodegeneration associated with Down's syndrome,
 CC neuropathic pain syndrome, Olivopontocerebral atrophy, amyotrophic
 CC lateral sclerosis, mitochondrial abnormalities, Alzheimer's disease,
 CC hepatic encephalopathy, Tourette's syndrome and drug addiction (all
 CC claimed). The lipophilic modified hedgehog polypeptide is also
 CC useful for promoting survival and/or functional performance of
 CC neuronal cells susceptible to excitotoxicity.

XX SQ Sequence 475 AA;
 Query Match 99.9%; Score 2467; DB 21; Length 475;
 Best Local Similarity 100.0%; Pred. No. 5.9e-218;
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLILARCLLVVSSLLVSGIACGPGRCFGRHRRHKKITPLAYKQIFINNAEKTIGASG 60
 DB 1 MLILARCLLVVSSLLVSGIACGPGRCFGRHRRHKKITPLAYKQIFINNAEKTIGASG 60
 QY 61 RYEGKISRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKINALAISVMNQPCV 120
 DB 61 RYEGKISRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKINALAISVMNQPCV 120
 QY 121 KLRVTEGMDDEGHHSEESIHYEGRAVDITSDRDSKCYMLARLAVEGFDVYTESKAH 180
 DB 121 KLRVTEGMDDEGHHSEESIHYEGRAVDITSDRDSKCYMLARLAVEGFDVYTESKAH 180
 QY 181 IHCYSVAENSVAKSGCGPFSATVHLEGGTKLVKDLSPGDRVLAADQGRILYSDFLT 240
 DB 181 IHCYSVAENSVAKSGCGPFSATVHLEGGTKLVKDLSPGDRVLAADQGRILYSDFLT 240
 QY 241 FLDRDDGAKKVVYIETREPRERLLITAAHLLFVAPHNDSATGEPBASSGSPSGGALG 300
 DB 241 FLDRDDGAKKVVYIETREPRERLLITAAHLLFVAPHNDSATGEPBASSGSPSGGALG 300
 QY 301 PRALFASRVPRQRYVYVVERDGRRLPLAAVHSYTLSEEAAGAVAPLTAOGTITILNRYL 360
 DB 301 PRALFASRVPRQRYVYVVERDGRRLPLAAVHSYTLSEEAAGAVAPLTAOGTITILNRYL 360
 QY 361 ASCYAVIEEHSNAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGA 420
 DB 361 ASCYAVIEEHSNAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGA 420
 QY 421 ADAPGAGATAGIHWYSOLLYOIGTWLIDSEALHPLGMAYKSSXSRCAGGAREGA 475
 DB 421 ADAPGAGATAGIHWYSOLLYOIGTWLIDSEALHPLGMAYKSSXSRCAGGAREGA 475

RESULT 8
 AAY96248
 ID AAY96248 standard; Protein; 475 AA.
 AC AAY96248;
 DT 11-SEP-2000 (first entry)
 DE Partial human Shh.
 KW Human; sonic hedgehog; Shh; neuromuscular disorder; neuropathy;
 KW Guillain-Barre syndrome; peripheral neuropathy; diabetes; alcoholism;
 KW chronic inflammatory demyelinating polyneuropathy; CIDP;
 KW gene therapy; infection; inflammation; hereditary neuropathy;
 KW Charcot-Marie-Tooth disease; vasculitis; lung cancer; tumour;
 KW multiple myeloma; nutritional imbalance; kidney disease;
 KW hypothyroid neuropathy; trauma; Refsum's disease; Abetalipoproteinemia;
 KW Tangle disease; Krabbe's disease; Metachromatic leukodystrophy;
 KW Fabry's disease; CMT; GBS; Dejerine-Sottas syndrome; acute neuropathy;
 KW amyotrophic lateral sclerosis; ALS; Miller-Fisher syndrome; amyloidosis;
 KW hereditary sensory neuropathy type II; HSN II; B-cell lymphoma;
 KW Waldenstrom's Macroglobulinemia; Chronic lymphocytic leukemia;
 KW neuroprotective; cytoprotective; patched-mediated signal transduction.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc-difference 463 /label= unknown
 FT /note= "Encoded by NNN"
 FT WO200027422-A2.

XX Sequence 475 AA: (99.9%) Score 2467; DB 21; Length 475;
 Query Match 100.0%; Pred. No. 5.9e-218;
 Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVYSSLLVCSGLACGPGFGKRRHPKLTPLAYKQFIPNVAEKTGASG 60
 DB 1 MLLARCLLVYSSLLVCSGLACGPGFGKRRHPKLTPLAYKQFIPNVAEKTGASG 60
 QY 61 RREGKISRNSERFKELTPYNDIIFKDEENTGADRLMTQCKDKINALAISVMNQPGV 120
 DB 61 RREGKISRNSERFKELTPYNDIIFKDEENTGADRLMTQCKDKINALAISVMNQPGV 120
 QY 121 KLRYTEGMDEGDHSESLHYEGRAVDITTSDBRSKYGMRLAYEAGFDWVYTESKAH 180
 DB 121 KLRYTEGMDEGDHSESLHYEGRAVDITTSDBRSKYGMRLAYEAGFDWVYTESKAH 180
 QY 181 IHCSYKAENSVAAKSGGCPGSATVHLEGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240
 DB 181 IHCSYKAENSVAAKSGGCPGSATVHLEGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240
 QY 241 FLDRODGAKKVYVYETREPRERLLTAHLLFVAPHNDSATGEPASSGSGPPSGGALG 300
 DB 241 FLDRODGAKKVYVYETREPRERLLTAHLLFVAPHNDSATGEPASSGSGPPSGGALG 300
 QY 301 PRALFASRYRPGORYVVAERDGRRLPAAVHSVTLSEAGAVAPLTAOGTILINRYL 360
 DB 301 PRALFASRYRPGORYVVAERDGRRLPAAVHSVTLSEAGAVAPLTAOGTILINRYL 360
 QY 361 ASCYAVIEHSHMAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGVALTAPGA 420
 DB 361 ASCYAVIEHSHMAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGVALTAPGA 420
 QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLDEALHPLGMAVKSXSRRGAGGAREGA 475
 DB 421 ADAPGAGATAGIHWYSQLLYQIGTWLDEALHPLGMAVKSXSRRGAGGAREGA 475

RESULT 10
 AAG65748
 ID AAG65748 standard; Protein; 475 AA.
 AC AAG65748;
 DT 07-JAN-2002 (first entry)
 DE Human sonic hedgehog (Shh) polypeptide.
 KW Adipocyte; hedgehog polypeptide; desert hedgehog; Indian hedgehog; Dhh;
 KW Ihh; sonic hedgehog; Shh; therapeutic; cytostatic; tumor.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc-difference 463 /note="encoded by NNN"
 XX WO200164238-A2.
 PD 07-SEP-2001.
 PF 28-FEB-2001; 2001MO-US06450.
 PR 29-FEB-2000; 2000US-186058P.
 PA (CURTI-) CURIS INC.
 XX Zehentner B, Leser-Reliff U, Bartscher H;
 XX WPI; 2001-607352/69.
 DR N-PSDB; AAI66776.

XX Method for regulating formation and/or maintenance of adipocyte tissue
 PT by contacting pre-adipocyte or adipocyte cells with a hedgehog
 PT polypeptide or plc therapeutic
 PS Disclosure; Page 102-104; 132pp; English.

XX The invention provides a method for regulating formation and/or
 CC maintenance of adipocyte tissue that comprises contacting pre adipocyte
 CC or adipocyte cells with a hedgehog polypeptide or plc therapeutic. The
 CC method is used for regulating the growth state of an adipocyte stem/
 CC progenitor cell, and treating or preventing disorders of, or surgical or
 CC cosmetic repair of, adipocyte tissues, e.g. for treating or preventing
 CC hyperplastic or neoplastic conditions affecting adipocyte tissue, such
 CC as soft tissue tumors, especially adipose cell tumors, e.g. lipomas,
 CC fibrolipomas, lipoblastomas, lipomatosis, hibernomas, hemangiomas and/or
 CC liposarcomas. Hedgehog polypeptides can be used in combination with other
 CC therapeutic agents. The present sequence represents a human sonic
 CC hedgehog (Shh) polypeptide.

XX Sequence 475 AA: (99.9%) Score 2467; DB 22; Length 475;
 Query Match 100.0%; Pred. No. 5.9e-218;
 Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVYSSLLVCSGLACGPGFGKRRHPKLTPLAYKQFIPNVAEKTGASG 60
 DB 1 MLLARCLLVYSSLLVCSGLACGPGFGKRRHPKLTPLAYKQFIPNVAEKTGASG 60
 QY 61 RREGKISRNSERFKELTPYNDIIFKDEENTGADRLMTQCKDKINALAISVMNQPGV 120
 DB 61 RREGKISRNSERFKELTPYNDIIFKDEENTGADRLMTQCKDKINALAISVMNQPGV 120
 QY 121 KLRYTEGMDEGDHSESLHYEGRAVDITTSDBRSKYGMRLAYEAGFDWVYTESKAH 180
 DB 121 KLRYTEGMDEGDHSESLHYEGRAVDITTSDBRSKYGMRLAYEAGFDWVYTESKAH 180
 QY 181 IHCSYKAENSVAAKSGGCPGSATVHLEGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240
 DB 181 IHCSYKAENSVAAKSGGCPGSATVHLEGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240
 QY 241 FLDRODGAKKVYVYETREPRERLLTAHLLFVAPHNDSATGEPASSGSGPPSGGALG 300
 DB 241 FLDRODGAKKVYVYETREPRERLLTAHLLFVAPHNDSATGEPASSGSGPPSGGALG 300
 QY 301 PRALFASRYRPGORYVVAERDGRRLPAAVHSVTLSEAGAVAPLTAOGTILINRYL 360
 DB 301 PRALFASRYRPGORYVVAERDGRRLPAAVHSVTLSEAGAVAPLTAOGTILINRYL 360
 QY 361 ASCYAVIEHSHMAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGVALTAPGA 420
 DB 361 ASCYAVIEHSHMAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGVALTAPGA 420
 QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLDEALHPLGMAVKSXSRRGAGGAREGA 475
 DB 421 ADAPGAGATAGIHWYSQLLYQIGTWLDEALHPLGMAVKSXSRRGAGGAREGA 475

RESULT 11
 AAB84674
 ID AAB84674 standard; Protein; 475 AA.
 AC AAB84674;
 DT 17-SEP-2001 (first entry)
 DE Amino acid sequence of a human hedgehog (Shh) polypeptide.
 KW Shh gene; hedgehog gene; T lymphocyte; patched gene; infection;
 KW diabetes; nutritional deficiency; graft rejection; hypercure response;
 KW cornea transplant; autoimmune disorder; multiple sclerosis; psoriasis;
 KW atopic dermatitis; inflammatory disease; proliferative disease;

KM hyperproliferative disease; eczematous dermatitis; urticaria;
 KW vasculitis; scleroderma.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc-difference 463
 TT /note= "unspecified residue encoded by NNN"
 PN MO200140438-A2.
 XX
 PD 07-JUN-2001.
 XX
 PF 30-NOV-2000; 2000MO-US32590.
 XX
 PR 30-NOV-1999; 99US-0168112.
 XX
 PA (CURT-) CURTIS INC.
 PI Crompton T;
 XX
 DR WPI: 2001-441484/47.
 N-PSDB; AAH28451.
 XX
 PT Modulating immune function comprises administration of a hedgehog or
 PT pc agent, for treating e.g. diabetes, eczematous dermatitis, urticaria
 PT or vasculitis -
 XX
 PS Claim 4; Page 84-86; 105pp; English.
 XX
 CC The present sequence represents a hedgehog (Shh) polypeptide. Hedgehog
 CC gene products and signal transduction pathways involving hedgehog are
 CC involved in the maturation of T lymphocytes. The specification describes
 CC a method for modulating immune function, by administration of a hedgehog
 CC or patched (ptc) polypeptide, agonists or antagonists. The method is
 CC used to treat disorders affecting the regulation of lymphocytes,
 CC particularly maturation and/or activation of T lymphocytes. It is used,
 CC to treat bacterial or viral infection, diabetes, nutritional
 CC deficiencies, graft rejection or other hyperacute response such as
 CC kidney, heart, lung, bone marrow spleen skin or cornea transplant or
 CC autoimmune disorders such as multiple sclerosis, psoriasis or atopic
 CC dermatitis. The method is used to treat inflammatory, proliferative and
 CC hyperproliferative diseases, as well as cutaneous manifestations of
 CC immunological disorders such as eczematous dermatitis, urticaria,
 CC vasculitis and scleroderma.
 CC
 SQ Sequence 475 AA;
 Query Match 99.98; Score 2467; DB 22; Length 475;
 Best Local Similarity 100.0%; Pred. No. 5.9e-218;
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 301 PRALEASRVPRGQRYVVAERDGRLLPAVHSVTLSSEAAAGVAPLTAOGTILINRYL 360
 QY 361 ASCYAVEESHMAHRAFPRLAALAPARTDGGSDGGGGRGGRVALTPAGA 420
 DB 361 ASCYAVEESHMAHRAFPRLAALAPARTDGGSDGGGGRGGRVALTPAGA 420
 QY 421 ADAPGAGATAGIMHYSQLYQIGTWLDSBALHPICNAVSSXSRSAGAGAREGA 475
 DB 421 ADAPGAGATAGIMHYSQLYQIGTWLDSBALHPICNAVSSXSRSAGAGAREGA 475
 RESULT 12
 AAED04687
 ID AAED04687 standard; Protein; 475 AA.
 XX
 AC AAED04687;
 XX
 DT 04-SEP-2001 (first entry)
 XX
 DE Human sonic hedgehog (Shh) protein.
 XX
 KW Human; hedgehog protein; noctropic; neuroprotective; anticonvulsant;
 KW cytoskeletal; therapy; Alzheimer's disease; Parkinson's disease; injury;
 KW Huntington's chorea; amyotrophic lateral sclerosis; multiple sclerosis;
 KW nervous system aging; neurodegenerative disease; immunological disease;
 KW malignant glioma; medulloblastoma; neuroectodermal tumour; cancer;
 KW extracellular signalling protein.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 24..197
 FT /note= "N-terminal fragment"
 FT Misc-difference 463
 FT /label= Unknown
 FT /note= "Encoded by NNN"
 PN MO200134654-A1.
 XX
 PD 17-MAY-2001.
 XX
 PF 02-NOV-2000; 2000MO-US30405.
 XX
 PR 05-NOV-1999; 99US-0164025.
 XX
 PA (BIOJ) BIOGEN INC.
 PI Strauch K;
 XX
 DR WPI: 2001-329075/34.
 N-PSDB; AAD09034.
 XX
 PT Novel isolated hedgehog fusion polypeptide useful for treating
 PT neurological conditions such as Alzheimer's disease, Parkinson's
 PT disease, Huntington's chorea, amyotrophic lateral sclerosis, and
 PT multiple sclerosis -
 XX
 PS Disclosure; Page 117-118; 178pp; English.
 XX
 CC The present invention relates to hedgehog fusion proteins. Hedgehog
 CC proteins are a family of extracellular signalling proteins that regulate
 CC various aspects of embryonic development both in vertebrates and in
 CC invertebrates. Hedgehog fusion protein is useful for the prophylaxis or
 CC treatment of any condition or disease state for which a hedgehog or
 CC patched protein constituent is efficacious and in the diagnosis of
 CC constituents and/or conditions of disease states in non-physiological system or
 CC specimens and for diagnostic purposes in non-physiological systems.
 CC Hedgehog fusion protein is useful for treating neurological conditions
 CC due to injury, aging of nervous system, including Alzheimer's disease,
 CC chronic neurodegenerative diseases of the nervous system, including
 CC Parkinson's disease, Huntington's chorea, amyotrophic lateral sclerosis
 CC and chronic immunological diseases of nervous system including multiple

CC sclerosis and malignant gliomas, medulloblastomas, neuroectodermal
CC tumors and to specifically target medical therapies against cancers and
CC tumors which express the receptor for the protein. The present sequence
CC is human sonic hedgehog (Shh) protein.

XX Sequence 475 AA:

Query Match 99.9%; Score 2467; DB 22; Length 475;
Best Local Similarity 100.0%; Pred. No. 5.9e-218;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSSLLVSGLAGCGRGFRGRRHKKITPLAYKQFIPNVAEKTIGASG 60
DB 1 MLLARCLLVVSSLLVSGLAGCGRGFRGRRHKKITPLAYKQFIPNVAEKTIGASG 60
QY 61 RYEGKISRNSERFEKLTPTNPNIDIFKDEENTGADRLMTORCKDKLNALATSVNMQPGV 120
DB 61 RYEGKISRNSERFEKLTPTNPNIDIFKDEENTGADRLMTORCKDKLNALATSVNMQPGV 120
QY 121 KLRVTEGMDGDHSESLHYEGRAVDITTSRDRSKYGMRLAVEAGFDWVYESKAH 180
DB 121 KLRVTEGMDGDHSESLHYEGRAVDITTSRDRSKYGMRLAVEAGFDWVYESKAH 180
QY 181 HICSVKAENSVAAKSGCGPGSATYHLBOGGTKLVKDLSPGDRVLAADQGLLYSDFLT 240
DB 181 HICSVKAENSVAAKSGCGPGSATYHLBOGGTKLVKDLSPGDRVLAADQGLLYSDFLT 240
QY 241 FLDRDGAKKVYVETREPRERLLTAAHLFVAPHNDSATGEPASSGSGPPSGALG 300
DB 241 FLDRDGAKKVYVETREPRERLLTAAHLFVAPHNDSATGEPASSGSGPPSGALG 300
QY 301 PALFASRVPRGQRYVVAERDGRRLPAVAHSTLTSEEAAGAVAPLTAOGTILINRYL 360
DB 301 PALFASRVPRGQRYVVAERDGRRLPAVAHSTLTSEEAAGAVAPLTAOGTILINRYL 360
QY 361 ASCYAVIEHSNAHRAFAFRLHALALAPARDRGDSCGGDGGGGRVALTAPAA 420
DB 361 ASCYAVIEHSNAHRAFAFRLHALALAPARDRGDSCGGDGGGGRVALTAPAA 420
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLIDSEALHPGLMAVKSXSRRAGGAREGA 475
DB 421 ADAPGAGATAGIHWSQLLYQIGTWLIDSEALHPGLMAVKSXSRRAGGAREGA 475

RESULT 13

AAE05377 standard; Protein; 475 AA.

XX AAE05377;

DT 12-SEP-2001 (first entry)

XX Human Sonic hedgehog protein.

KM Human; Sonic hedgehog; Shh; morphogenic signal; neuron; chromosome 7q;

KM embryonic patterning; cell culture; cell differentiation; ischaemia;

KM cell proliferative disorder; intracerebral grafting; Huntington's chorea;

KM neurological disorder; Alzheimer's disease; Parkinson's disease;

XX amyotrophic lateral sclerosis; ALS; multiple sclerosis.

OS Homo sapiens.

XX Key

XX Misc-difference 463

XX 17-JUL-2001.

XX 02-JUL-1996; 96US-0674509.

PR 30-DEC-1993; 93US-0176427.
PR 14-DEC-1994; 94US-0356060.
PR 04-MAY-1995; 95US-0435093.
PR 05-JUN-1995; 95US-0460900.
PR 05-JUN-1995; 95US-0462386.

XX (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.

PA (HARD) HARVARD COLLEGE.

XX Marigo V, Tablin CJ, Ingham PW, McMahon AP;

XX WPI, 2001-440859/47.

XX N-PSDB: AAD10151.

PT Screening compounds that potentiate or inhibit binding of hedgehog
PT polypeptide to naturally occurring patched receptor, comprises
PT contacting polypeptide with receptor and test compound, and detecting
PT change in binding

PS Claim 2; Column 163-166; 127pp; English.

CC The present invention relates to assay for screening compounds that
CC potentiate or inhibit binding of hedgehog polypeptide to naturally
CC occurring patched receptor. The hedgehog proteins comprise morphogenic
CC signals produced by embryonic patterning centres, and are involved in the
CC formation and maintenance of ordered spatial arrangements of
CC differentiated tissues in vertebrates, both adult and embryonic. The
CC proteins can be used to generate and/or maintain an array of different
CC vertebrate tissues both in vitro and in vivo. The invention also relates
CC to a method for modulating growth, differentiation or survival of a
CC mammalian cell (e.g. neuron, testicular cell) responsive to hedgehog
CC induction. Hedgehog agonists and antagonists can be used in cell culture
CC techniques to enhance survival and maintenance of neurons and various
CC vertebrate organogenic pathways. The hedgehog gene is useful in
CC determining whether a patient is at the risk of disorder characterised by
CC unwanted cell proliferation or aberrant control of differentiation. The
CC hedgehog proteins or mimetics can be used to induce foetal neurons
CC especially neuronal stem cells in intracerebral grafting. The protein
CC or its mimetic can be used in the treatment of neurological conditions
CC e.g. injury to nervous system, ischaemia resulting from stroke,
CC Alzheimer's disease, Parkinson's disease, Huntington's chorea,
CC amyotrophic lateral sclerosis (ALS) and multiple sclerosis. The present
CC sequence is human Sonic hedgehog (Shh) protein. The Shh gene is
CC located on human chromosome 7q.

XX Sequence 475 AA:

Query Match 99.9%; Score 2467; DB 22; Length 475;
Best Local Similarity 100.0%; Pred. No. 5.9e-218;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSSLLVSGLAGCGRGFRGRRHKKITPLAYKQFIPNVAEKTIGASG 60
DB 1 MLLARCLLVVSSLLVSGLAGCGRGFRGRRHKKITPLAYKQFIPNVAEKTIGASG 60
QY 61 RYEGKISRNSERFEKLTPTNPNIDIFKDEENTGADRLMTORCKDKLNALATSVNMQPGV 120
DB 61 RYEGKISRNSERFEKLTPTNPNIDIFKDEENTGADRLMTORCKDKLNALATSVNMQPGV 120
QY 121 KLRVTEGMDGDHSESLHYEGRAVDITTSRDRSKYGMRLAVEAGFDWVYESKAH 180
DB 121 KLRVTEGMDGDHSESLHYEGRAVDITTSRDRSKYGMRLAVEAGFDWVYESKAH 180
QY 181 HICSVKAENSVAAKSGCGPGSATYHLBOGGTKLVKDLSPGDRVLAADQGLLYSDFLT 240
DB 181 HICSVKAENSVAAKSGCGPGSATYHLBOGGTKLVKDLSPGDRVLAADQGLLYSDFLT 240
QY 241 FLDRDGAKKVYVETREPRERLLTAAHLFVAPHNDSATGEPASSGSGPPSGALG 300
DB 241 FLDRDGAKKVYVETREPRERLLTAAHLFVAPHNDSATGEPASSGSGPPSGALG 300
QY 301 PALFASRVPRGQRYVVAERDGRRLPAVAHSTLTSEEAAGAVAPLTAOGTILINRYL 360

Db 301 PRLAFASRVPRGORYVVAERDGRLLPAAVHSVTISEBAAGYAPLTQGTILINRVL 360
 QY 361 ASCYAVIEESHMAHRAFPRLAHALLAALAPARTDRGDSGGDRGGCGRYALTAPGA 420
 Db 361 ASCYAVIEESHMAHRAFPRLAHALLAALAPARTDRGDSGGDRGGCGRYALTAPGA 420
 QY 421 ADAPGAGATGIMHYSQLLYQIGTWLDSBALHPLGMAVSSXSRCAGGARREGA 475
 Db 421 ADAPGAGATGIMHYSQLLYQIGTWLDSBALHPLGMAVSSXSRCAGGARREGA 475
 RESULT 14
 ID AAB31222 standard; Protein: 475 AA.
 AC AAB31222;
 XX 20-APR-2001 (first entry)
 DE Amino acid sequence of human sonic hedgehog protein (Shh).
 XX Hedgehog related-protein; sonic hedgehog protein; Shh; Ischemia; stroke;
 KM desert hedgehog protein; Dhh; Indian hedgehog protein; Ihh; neuron;
 KM neurological condition; nervous system injury; tumour-induced injury;
 KM aging; Alzheimer's disease; chronic neurodegenerative disease;
 KM Parkinson's disease; Huntington's chorea; amyotrophic lateral sclerosis;
 KM spinocerebellar degeneration; chronic immunological disease;
 XX multiple sclerosis.
 XX Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc-difference 463
 FT /note="unspecified amino acid encoded by NNN"
 XX US6155747-A.
 PD 26-DEC-2000.
 PF 05-JUN-1995; 95US-0460900.
 XX 30-DEC-1993; 93US-0176427.
 PR 14-DEC-1994; 94US-0356060.
 PR 04-MAY-1995; 95US-0435093.
 XX
 PA (HARD) HARVARD COLLEGE.
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
 PI Ingham PW, McMahon AB, Tablin CJ, Marti-gorostiza E, Bumcrot DA;
 XX N-PSDB; AAC87079.
 DR WPI: 2001-079847/09.
 DR N-PSDB; AAC87079.
 PT Polynucleotides encoding hedgehog proteins, useful for treating
 PT diseases of nervous system such as Alzheimer's disease, Parkinson's
 PT disease, Huntington's chorea, amyotrophic lateral sclerosis, multiple
 PT sclerosis -
 PS Claim 10; Columns 147-150; 11pp; English.
 XX The present sequence represents a hedgehog related-protein. The
 CC specification describes a sonic hedgehog protein (Shh), a desert
 CC hedgehog protein (Dhh), and an Indian hedgehog protein (Ihh). The
 CC hedgehog polynucleotides are useful in diagnostic, in antisense
 CC therapy and in therapeutic assays for detecting and treating disorders
 CC involving, e.g., aberrant expression of vertebrate hedgehog homologue.
 CC hedgehog polypeptides are useful therapeutically to enhance survival
 CC of neurons and other neuron cells and in treating neurological
 CC conditions deriving from acute, subacute, or chronic injury to the
 CC nervous system, including traumatic injury, chemical injury, vascular
 CC injury and deficits (such as the ischemia resulting from stroke),
 CC together with infectious/inflammatory and induced-induced injury, aging
 CC of the nervous system including Alzheimer's disease, chronic

CC neurodegenerative diseases of the nervous system, including Parkinson's
 CC disease, Huntington's chorea, amyotrophic lateral sclerosis,
 CC spinocerebellar degenerations, and chronic immunological diseases of
 CC the nervous system or affecting the nervous system, including multiple
 CC sclerosis.
 XX Sequence 475 AA:
 SQ
 Query Match 99.9%; Score 2467; DB 22; Length 475;
 Best Local Similarity 100.0%; Pred. No. 5,9e-218;
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLLIARCLLLVLYSSLYVSGLAGCGPGRGFKRRHPKRLPLAYKQFIPVNAEKTIGASG 60
 Db 1 MLLIARCLLLVLYSSLYVSGLAGCGPGRGFKRRHPKRLPLAYKQFIPVNAEKTIGASG 60
 QY 61 RYEGKISRNSERREKELPNPNPDIFKDEENTGADRLMTQCKDKNALISVNMOPGV 120
 Db 61 RYEGKISRNSERREKELPNPNPDIFKDEENTGADRLMTQCKDKNALISVNMOPGV 120
 QY 121 KLRYTEGMDGDGHHSESLHYEGRAVDITTSDBDRSKYGMILARLAYAGPDWYVESKAH 180
 Db 121 KLRYTEGMDGDGHHSESLHYEGRAVDITTSDBDRSKYGMILARLAYAGPDWYVESKAH 180
 QY 181 IHCYKAEANSVAAKSGGCFPGSATVHLEGGCTKLVKDLSFGDRYLADDDGRLLYSDFLT 240
 Db 181 IHCYKAEANSVAAKSGGCFPGSATVHLEGGCTKLVKDLSFGDRYLADDDGRLLYSDFLT 240
 QY 241 FLDRDCAKKVFYIERERERERLLTAHLILVPAHPNDATGEPKSSSGSPSGCALG 300
 Db 241 FLDRDCAKKVFYIERERERERLLTAHLILVPAHPNDATGEPKSSSGSPSGCALG 300
 QY 301 PRLAFASRVPRGORYVVAERDGRLLPAAVHSVTISEBAAGYAPLTQGTILINRVL 360
 Db 301 PRLAFASRVPRGORYVVAERDGRLLPAAVHSVTISEBAAGYAPLTQGTILINRVL 360
 QY 361 ASCYAVIEESHMAHRAFPRLAHALLAALAPARTDRGDSGGDRGGCGRYALTAPGA 420
 Db 361 ASCYAVIEESHMAHRAFPRLAHALLAALAPARTDRGDSGGDRGGCGRYALTAPGA 420
 QY 421 ADAPGAGATGIMHYSQLLYQIGTWLDSBALHPLGMAVSSXSRCAGGARREGA 475
 Db 421 ADAPGAGATGIMHYSQLLYQIGTWLDSBALHPLGMAVSSXSRCAGGARREGA 475
 RESULT 15
 ID AAB60265 standard; Protein: 475 AA.
 AC AAB60265;
 XX 30-MAR-2001 (first entry)
 DE Human Sonic hedgehog (Shh) protein, SEQ ID NO:15.
 XX Hedgehog protein; polymer conjugate; polyalkene glycol group;
 KM bioavailability; formulation; neurological disorder;
 KM inflammatory disorder; autoimmune disorder; cancer;
 KM neurodegenerative disorder; Parkinson's disease; Huntington's disease;
 KM Alzheimer's disease; neurological injury; stroke; multiple sclerosis;
 KM malignant glioma; medulloblastoma; neuroectodermal tumour.
 OS Homo sapiens.
 PN WO200073337-A1.
 PD 07-DEC-2000.
 PF 26-MAY-2000; 2000WO-US14741.
 PR 01-JUN-1999; 99US-0137011.
 PR 13-AUG-1999; 99US-0149016.
 XX

PA (BIOD) BIOGEN INC.
XX
XX Pepinsky RB, Taylor F, Garber E;
XX
XX WPI: 2001-049927/06.
DR N-PSDB: AAF27018.
XX
XX Modified hedgehog protein, useful in the treatment of Parkinson's
PT disease and Huntington's chorea, comprises a polymer containing a
PT polyalkylene glycol group linked to any residue other than the
PT N-terminal and lysine residues -
XX
XX
PS Disclosure: Page 138-140; 157pp; English.
XX
XX The invention relates to novel polymer conjugates of hedgehog proteins
CC which have increased bioavailability. The hedgehog proteins are
CC conjugated to a non-naturally-occurring polymer comprising a polyalkylene
CC glycol group, with the proviso that the polymer is not conjugated to the
CC N-terminus, or to lysine residues of the hedgehog protein. The hedgehog
CC protein used in the conjugate may be a wild-type or mutant Sonic hedgehog
CC (Shh), Indian hedgehog (Ihh) or Desert hedgehog (Dhh) protein, or may be
CC a hedgehog fusion protein. The invention also relates to methods of
CC defining and mapping functionally important regions of a protein by
CC modifying accessible amino acid side chains, and determining the effect
CC the position and/or type of modification have on the activity of the
CC protein. The hedgehog polymer conjugates may be used in the management of
CC various medical conditions including various neurological disorders,
CC inflammatory and autoimmune diseases, and cancers. In particular, they
CC may be used to prevent preventing or ameliorate neurodegenerative
CC disorders (e.g., Parkinson's disease, Huntington's disease, Alzheimer's
CC disease); age-associated neurological disease; neurological injury and
CC trauma; immunological diseases of the nervous system (e.g., multiple
CC sclerosis); stroke; and malignant gliomas, medulloblastomas and
CC neuroectodermal tumours. The modifications made to the hedgehog protein
CC may result in increased half-life, altered tissue distribution (such as
CC an improved ability to stay in the vasculature for longer periods of
CC time), increased stability in solution, protection from proteolytic
CC degradation, or reduced immunogenicity. In particular, the ability to
CC remain in the vasculature for prolonged periods may allow a hedgehog
CC protein of the invention to cross the blood-brain barrier, and an
CC increased thermal stability would be an advantage when formulating the
CC hedgehog protein in powder form. The present sequence represents a
CC member of the hedgehog family of proteins.
XX
XX

SQ Sequence 475 AA:

Query Match 99.9%; Score 2467; DB 22; Length 475;
Best Local Similarity 100.0%; Pred. No. 5.9e-218;

Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLLARCLLVVSSLLVCSGLACGPGRGFRHPRKLTPLAYKQFIIPNVAEKTLAGSG 60
DB 1 MLLARCLLVVSSLLVCSGLACGPGRGFRHPRKLTPLAYKQFIIPNVAEKTLAGSG 60
OY 61 RYEGKISRSEFEKELTPYNNPDIIFKDEENTGADRLMTQRCCKLNLALISVNMQPGV 120
DB 61 RYEGKISRSEFEKELTPYNNPDIIFKDEENTGADRLMTQRCCKLNLALISVNMQPGV 120
OY 121 KLRVTEGMDDEHHSEESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYYSKAH 180
DB 121 KLRVTEGMDDEHHSEESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYYSKAH 180
OY 181 IHCYKAVNSVAAKSGGCPGSAIVHLERGGTKLVKDLSPGDRVLAADQGRILYSDFLT 240
DB 181 IHCYKAVNSVAAKSGGCPGSAIVHLERGGTKLVKDLSPGDRVLAADQGRILYSDFLT 240
OY 241 FLDRDDGAKKVFYVETREPRERLLTAHLFLVAPHNDSATGEPEASSGSGPPSGALG 300
DB 241 FLDRDDGAKKVFYVETREPRERLLTAHLFLVAPHNDSATGEPEASSGSGPPSGALG 300
OY 301 PALAFASRYRPGORYVVAERDGRRLPAVHSTLSEEAAGAYAPLTAOCTILINRYL 360
DB 301 PALAFASRYRPGORYVVAERDGRRLPAVHSTLSEEAAGAYAPLTAOCTILINRYL 360

OY 361 ASCYAVIEEHSNAHRAFAFPRLAHALLALAPARTDRCGSGGDRGGGVALTAPGA 420
DB 361 ASCYAVIEEHSNAHRAFAFPRLAHALLALAPARTDRCGSGGDRGGGVALTAPGA 420
OY 421 ADAPGAGATAGIHWYSQLLYQIGTWLIDSEALHPICMAVKSXSRSRAGCGAREGA 475
DB 421 ADAPGAGATAGIHWYSQLLYQIGTWLIDSEALHPICMAVKSXSRSRAGCGAREGA 475

Search completed: February 20, 2003, 10:10:20
Job time : 36.4615 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:08:38 ; Search time 12.4346 Seconds
(Without alignments)
1123.956 Million cell updates/sec

Title: US-09-827-110A-15

Perfect score: 2469
Sequence: 1 MLLARCLLVVSLVCS.....GMAVKSXRGAGGAREGA 475

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2467	99.9	475	2	US-08-356-060A-13
2	2467	99.9	475	4	US-08-460-900C-13
3	2467	99.9	475	4	US-08-674-509B-13
4	2467	99.9	475	4	US-08-954-698-13
5	2467	99.9	475	4	US-08-954-698-13
6	2467	99.9	475	4	US-08-957-874-13
7	2407	97.5	462	1	US-09-325-256-22
8	2407	97.5	462	1	US-08-748-591-4
9	2068	83.8	437	3	US-08-946-329A-20
10	2068	83.8	437	4	US-08-567-357A-20
11	2068	83.8	437	4	US-08-729-743A-20
12	2068	83.8	437	4	US-09-057-860A-6
13	2068	83.8	437	4	US-08-349-498-20
14	2068	83.8	437	5	PCT-US95-15463-20
15	2068	83.8	437	5	PCT-US95-15463-20
16	2065	83.6	437	1	US-08-176-427B-8
17	2065	83.6	437	2	US-08-356-060A-11
18	2065	83.6	437	4	US-08-460-900C-11
19	2065	83.6	437	4	US-08-674-509B-11
20	2065	83.6	437	4	US-08-954-698-11
21	2065	83.6	437	4	US-08-957-874-11
22	2065	83.6	437	4	US-09-325-256-20
23	2062	83.5	437	4	US-09-293-505-14
24	2060	83.4	437	4	US-08-757-230A-2
25	2060	83.4	437	4	US-08-757-230A-2
26	2060	83.4	437	5	PCT-US95-02315-2
27	1807	73.2	425	1	US-08-176-427B-2

28	1807	73.2	425	2	US-08-356-060A-8	Sequence 8, Appl1
29	1807	73.2	425	4	US-08-460-900C-8	Sequence 8, Appl1
30	1807	73.2	425	4	US-08-674-509B-8	Sequence 8, Appl1
31	1807	73.2	425	4	US-08-954-698-8	Sequence 8, Appl1
32	1807	73.2	425	4	US-08-954-698-8	Sequence 8, Appl1
33	1807	73.2	425	4	US-09-325-256-17	Sequence 17, Appl1
34	1690	68.4	425	3	US-08-946-329A-19	Sequence 19, Appl1
35	1690	68.4	425	4	US-08-567-357A-19	Sequence 19, Appl1
36	1690	68.4	425	4	US-08-729-743A-19	Sequence 19, Appl1
37	1690	68.4	425	4	US-08-349-498-19	Sequence 19, Appl1
38	1690	68.4	425	5	PCT-US95-15463-19	Sequence 19, Appl1
39	1690	68.4	425	5	PCT-US95-15463-19	Sequence 19, Appl1
40	1506	61.0	418	4	US-08-946-329A-18	Sequence 18, Appl1
41	1506	61.0	418	4	US-08-567-357A-18	Sequence 18, Appl1
42	1506	61.0	418	4	US-08-729-743A-18	Sequence 18, Appl1
43	1506	61.0	418	4	US-08-757-230A-7	Sequence 7, Appl1
44	1506	61.0	418	4	US-08-349-498-18	Sequence 18, Appl1
45	1506	61.0	418	5	PCT-US95-15463-18	Sequence 18, Appl1

ALIGNMENTS

RESULT 1
US-08-356-060A-13
Sequence 13, Application US/08356060A
Patent No. 5844079

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,060A
FILING DATE: 14-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006CP
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 475 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-060A-13

Query Match 99.9%; Score 2467; DB 2; Length 475;
Best Local Similarity 100.0%; Pred. No. 8, 2e-239;

Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSLVCSGLACGPGRGKRRHKKLTPLAYKQFIPNVAERTKLGASC 60

|||||
Db 1 MILLARCLLVVSSLLVSGLACGPGFGRRHPKKTLPYAKOFTIPNVAEKTGASG 60
QY 61 RYEGKISRNSERFKELTPYNDPIIFKDEENTGADRLMTORCKDKLNALAI SVMNMPGV 120
Db 61 RYEGKISRNSERFKELTPYNDPIIFKDEENTGADRLMTORCKDKLNALAI SVMNMPGV 120
QY 121 KLRVTEGMDECHHSEESIHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYESKAH 180
Db 121 KLRVTEGMDECHHSEESIHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYESKAH 180
QY 181 HCSYKAENSVAKSGGCGPGSATVHLEOGGTRKYKDLSPGSRVLAADOGRLTSDFLT 240
Db 181 HCSYKAENSVAKSGGCGPGSATVHLEOGGTRKYKDLSPGSRVLAADOGRLTSDFLT 240
QY 241 FLDRDGAKKVYVETREPRERLLTAAHLTFVAPHNDATGEPEASSGSGPPSGALG 300
Db 241 FLDRDGAKKVYVETREPRERLLTAAHLTFVAPHNDATGEPEASSGSGPPSGALG 300
QY 301 PRALFASRVPRGQRYVVAERDGRRLPAAVHSTVLSSEAAGAVAPLTAOCTIILINRVL 360
Db 301 PRALFASRVPRGQRYVVAERDGRRLPAAVHSTVLSSEAAGAVAPLTAOCTIILINRVL 360
QY 361 ASCYAVIEESHNAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGVALTAPGA 420
Db 361 ASCYAVIEESHNAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGVALTAPGA 420
QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLDSSEALHPGLMAVKSXSXSGAGGAREGA 475
Db 421 ADAPGAGATAGIHWYSQLLYQIGTWLDSSEALHPGLMAVKSXSXSGAGGAREGA 475

RESULT 2
US-08-460-900C-13
; Sequence 13, Application US/08460900C
; Patent No. 6165747
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tablin, Clifford J.
; APPLICANT: Bumcrot, David A.
; APPLICANT: Morli-Gorositz, Elisa
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; TITLE OF INVENTION: Proteins and Uses Related Thereto
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,900C
; FILING DATE: 5-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/435,093
; FILING DATE: 4-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/356,060
; FILING DATE: 14-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HWY-006.05

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-460-900C-13

Query Match 99.9%; Score 2467; DB 4; Length 475;
Best Local Similarity 100.0%; Pred. No. 8,2e-239;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|||||
QY 1 MILLARCLLVVSSLLVSGLACGPGFGRRHPKKTLPYAKOFTIPNVAEKTGASG 60
Db 1 MILLARCLLVVSSLLVSGLACGPGFGRRHPKKTLPYAKOFTIPNVAEKTGASG 60
QY 61 RYEGKISRNSERFKELTPYNDPIIFKDEENTGADRLMTORCKDKLNALAI SVMNMPGV 120
Db 61 RYEGKISRNSERFKELTPYNDPIIFKDEENTGADRLMTORCKDKLNALAI SVMNMPGV 120
QY 121 KLRVTEGMDECHHSEESIHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYESKAH 180
Db 121 KLRVTEGMDECHHSEESIHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYESKAH 180
QY 181 HCSYKAENSVAKSGGCGPGSATVHLEOGGTRKYKDLSPGSRVLAADOGRLTSDFLT 240
Db 181 HCSYKAENSVAKSGGCGPGSATVHLEOGGTRKYKDLSPGSRVLAADOGRLTSDFLT 240
QY 241 FLDRDGAKKVYVETREPRERLLTAAHLTFVAPHNDATGEPEASSGSGPPSGALG 300
Db 241 FLDRDGAKKVYVETREPRERLLTAAHLTFVAPHNDATGEPEASSGSGPPSGALG 300
QY 301 PRALFASRVPRGQRYVVAERDGRRLPAAVHSTVLSSEAAGAVAPLTAOCTIILINRVL 360
Db 301 PRALFASRVPRGQRYVVAERDGRRLPAAVHSTVLSSEAAGAVAPLTAOCTIILINRVL 360
QY 361 ASCYAVIEESHNAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGVALTAPGA 420
Db 361 ASCYAVIEESHNAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGVALTAPGA 420
QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLDSSEALHPGLMAVKSXSXSGAGGAREGA 475
Db 421 ADAPGAGATAGIHWYSQLLYQIGTWLDSSEALHPGLMAVKSXSXSGAGGAREGA 475

RESULT 3
US-08-674-509B-13
; Sequence 13, Application US/08674509B
; Patent No. 6261786
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tablin, Clifford J.
; APPLICANT: Marigo, Valeria
; TITLE OF INVENTION: SCREENING ASSAYS FOR HEDGEGOG AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/674,509B
 FILING DATE: 02-JUL-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/460,900
 FILING DATE: 05-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: HMV-006,06
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 475 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-674-509B-13

Query Match 99.9% Score 2467; DB 4; Length 475;
 Best Local Similarity 100.0%; Pred. No. 8.2e-239;
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFGKRRHPKKLPPLAYKQFIPNVAEKTIGASG 60
 DB 1 MLLARCLLVVSSLLVCSGLACGPGRGFGKRRHPKKLPPLAYKQFIPNVAEKTIGASG 60
 QY 61 RYEGKISNSERFELPNYNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMQMPGV 120
 DB 61 RYEGKISNSERFELPNYNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMQMPGV 120
 QY 121 KLRTYEGWDEGHHSEESLHTEGRAVDITTSDRRSKYGMRLAVEGFPWVYESSAH 180
 DB 121 KLRTYEGWDEGHHSEESLHTEGRAVDITTSDRRSKYGMRLAVEGFPWVYESSAH 180
 QY 181 IHCSVKAENSVAAKSGCGFPGSATVHLEQGSKLVKDLSPGDRVLAADDGRLLYSDFLT 240
 DB 181 IHCSVKAENSVAAKSGCGFPGSATVHLEQGSKLVKDLSPGDRVLAADDGRLLYSDFLT 240
 QY 241 FLRDDGAKKVFYIETREPERLLTFAHLLFVAPHNDSATGEPEASSSGSPSGALG 300
 DB 241 FLRDDGAKKVFYIETREPERLLTFAHLLFVAPHNDSATGEPEASSSGSPSGALG 300
 QY 301 PRALEASRVRRGORYVVAERDGRRLPAVHSTVLTSEEAAGVAPLTAAGTILINRVL 360
 DB 301 PRALEASRVRRGORYVVAERDGRRLPAVHSTVLTSEEAAGVAPLTAAGTILINRVL 360
 QY 361 ASCYAVIEHSMARAFAPFLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420
 DB 361 ASCYAVIEHSMARAFAPFLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420
 QY 421 ADAPGAGTAGIHWYSOLLIOIGTWLDSALHPLGMAVKSXSXRGAGGARREGA 475
 DB 421 ADAPGAGTAGIHWYSOLLIOIGTWLDSALHPLGMAVKSXSXRGAGGARREGA 475

RESULT 4
 US-08-954-698-13
 Sequence 13, Application US/08954698
 Patent No. 6271363

GENERAL INFORMATION:
 APPLICANT: Ingham, Phillip W.
 APPLICANT: McMahon, Andrew P.
 APPLICANT: Tablin, Clifford J.
 TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
 NUMBER OF SEQUENCES: 48
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
 STREET: One Post Office Square
 CITY: Boston

STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/954,698
 FILING DATE: 20-OCT-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/462,386
 FILING DATE: 05-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/435,093
 FILING DATE: 04-MAY-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/356,060
 FILING DATE: 14-DEC-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/176,427
 FILING DATE: 30-DEC-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: HMV-006,10
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 475 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-954-698-13

Query Match 99.9% Score 2467; DB 4; Length 475;
 Best Local Similarity 100.0%; Pred. No. 8.2e-239;
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFGKRRHPKKLPPLAYKQFIPNVAEKTIGASG 60
 DB 1 MLLARCLLVVSSLLVCSGLACGPGRGFGKRRHPKKLPPLAYKQFIPNVAEKTIGASG 60
 QY 61 RYEGKISNSERFELPNYNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMQMPGV 120
 DB 61 RYEGKISNSERFELPNYNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMQMPGV 120
 QY 121 KLRTYEGWDEGHHSEESLHTEGRAVDITTSDRRSKYGMRLAVEGFPWVYESSAH 180
 DB 121 KLRTYEGWDEGHHSEESLHTEGRAVDITTSDRRSKYGMRLAVEGFPWVYESSAH 180
 QY 181 IHCSVKAENSVAAKSGCGFPGSATVHLEQGSKLVKDLSPGDRVLAADDGRLLYSDFLT 240
 DB 181 IHCSVKAENSVAAKSGCGFPGSATVHLEQGSKLVKDLSPGDRVLAADDGRLLYSDFLT 240
 QY 241 FLRDDGAKKVFYIETREPERLLTFAHLLFVAPHNDSATGEPEASSSGSPSGALG 300
 DB 241 FLRDDGAKKVFYIETREPERLLTFAHLLFVAPHNDSATGEPEASSSGSPSGALG 300
 QY 301 PRALEASRVRRGORYVVAERDGRRLPAVHSTVLTSEEAAGVAPLTAAGTILINRVL 360
 DB 301 PRALEASRVRRGORYVVAERDGRRLPAVHSTVLTSEEAAGVAPLTAAGTILINRVL 360
 QY 361 ASCYAVIEHSMARAFAPFLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420
 DB 361 ASCYAVIEHSMARAFAPFLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420
 QY 421 ADAPGAGTAGIHWYSOLLIOIGTWLDSALHPLGMAVKSXSXRGAGGARREGA 475
 DB 421 ADAPGAGTAGIHWYSOLLIOIGTWLDSALHPLGMAVKSXSXRGAGGARREGA 475

RESULT 5
US-08-957-874-13

Sequence 13, Application US/08957874
Patent No. 6384192

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/957,874
FILING DATE: 20-OCT-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/462,386
FILING DATE: 5-JUNE-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/435,093
FILING DATE: 4-MAY-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/356,060
FILING DATE: 14-DEC-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:
LENGTH: 475 amino acids

TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-957-874-13

Query Match 99.9%; Score 2467; DB 4; Length 475;
Best Local Similarity 100.0%; Pred. No. 8.2e-239;

Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSSLLVCSGLACGPGFGKRRHKITPLAYKQFIPNVAEKTIGASG 60
DB 1 MLLARCLLVVSSLLVCSGLACGPGFGKRRHKITPLAYKQFIPNVAEKTIGASG 60
QY 61 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAISVMNQMPGV 120
DB 61 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAISVMNQMPGV 120
QY 121 KLRVTEGMDEDEHSESLHYEGRAVDITTSRDRSKYGMLARLAVEAGFDVYVESKAH 180
DB 121 KLRVTEGMDEDEHSESLHYEGRAVDITTSRDRSKYGMLARLAVEAGFDVYVESKAH 180
QY 181 IHCYKAENSVAAKSGGCPGSAIVHLEGGTKLVKDLSPGDRVLAADQGRILLYSDFLT 240
DB 181 IHCYKAENSVAAKSGGCPGSAIVHLEGGTKLVKDLSPGDRVLAADQGRILLYSDFLT 240

DB 181 IHCYKAENSVAAKSGGCPGSAIVHLEGGTKLVKDLSPGDRVLAADQGRILLYSDFLT 240
QY 241 FLDRDDGAKKVVYIETREPRERILLTAHLLFVAPHNDSATGEPPASSGPPSGGALG 300
DB 241 FLDRDDGAKKVVYIETREPRERILLTAHLLFVAPHNDSATGEPPASSGPPSGGALG 300
QY 301 PRALFASRVPGQRYVVAERDGRRLPAVAHSVTLSEEAAGAVAPLTAOGTILINRYL 360
DB 301 PRALFASRVPGQRYVVAERDGRRLPAVAHSVTLSEEAAGAVAPLTAOGTILINRYL 360
QY 361 ASCYAVIEESHNAHRAFAFRLAHALLALAPARDRGDSSGGDRGGGGVATAPGA 420
DB 361 ASCYAVIEESHNAHRAFAFRLAHALLALAPARDRGDSSGGDRGGGGVATAPGA 420
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLIDSEALHPLGMAYKSSXSRGAGGAREGA 475
DB 421 ADAPGAGATAGIHWSQLLYQIGTWLIDSEALHPLGMAYKSSXSRGAGGAREGA 475

RESULT 6
US-09-325-256-22

Sequence 22, Application US/09325256
Patent No. 6444793

GENERAL INFORMATION:

APPLICANT: PEPINSKY, R. BLAKE
APPLICANT: BAKER, DARREN P.
APPLICANT: WEN, DINGYI
APPLICANT: WILLIAMS, KEVIN P.
APPLICANT: GARGER, ELLEN A.
APPLICANT: TAYLOR, FREDERICK R.
APPLICANT: GALDES, ALPHONSE
APPLICANT: PORTER, JEFFREY

TITLE OF INVENTION: HYDROPHOBICALLY-MODIFIED PROTEIN COMPOSITIONS AND
TITLE OF INVENTION: METHODS

FILE REFERENCE: BY-067.01

CURRENT APPLICATION NUMBER: US/09/325,256
CURRENT FILING DATE: 1999-06-03

PRIOR APPLICATION NUMBER: 60/099,800
PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/078,935
PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/089,685
PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: 60/067,423
PRIOR FILING DATE: 1997-12-03

PRIOR APPLICATION NUMBER: PCT/US98/25676
PRIOR FILING DATE: 1998-12-03

NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 22
LENGTH: 475

TYPE: PRT
ORGANISM: Homo sapiens

FEATURE:
OTHER INFORMATION: Xaa may be any amino acid

US-09-325-256-22

Query Match 99.9%; Score 2467; DB 4; Length 475;
Best Local Similarity 100.0%; Pred. No. 8.2e-239;

Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSSLLVCSGLACGPGFGKRRHKITPLAYKQFIPNVAEKTIGASG 60
DB 1 MLLARCLLVVSSLLVCSGLACGPGFGKRRHKITPLAYKQFIPNVAEKTIGASG 60
QY 61 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAISVMNQMPGV 120
DB 61 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAISVMNQMPGV 120
QY 121 KLRVTEGMDEDEHSESLHYEGRAVDITTSRDRSKYGMLARLAVEAGFDVYVESKAH 180
DB 121 KLRVTEGMDEDEHSESLHYEGRAVDITTSRDRSKYGMLARLAVEAGFDVYVESKAH 180
QY 181 IHCYKAENSVAAKSGGCPGSAIVHLEGGTKLVKDLSPGDRVLAADQGRILLYSDFLT 240
DB 181 IHCYKAENSVAAKSGGCPGSAIVHLEGGTKLVKDLSPGDRVLAADQGRILLYSDFLT 240

QY 181 IHCYKAKENSVAAKSGCCFGPSATVHLEOGCTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240
DB 181 IHCYKAKENSVAAKSGCCFGPSATVHLEOGCTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240
QY 241 FLDDDDAKKVFYIETREPERERLLTFAHLLFPAPHNDSATGEPKSSGSPSGALG 300
DB 241 FLDDDDAKKVFYIETREPERERLLTFAHLLFPAPHNDSATGEPKSSGSPSGALG 300
QY 301 PRALFASVRPGQVYVVAERDGRLLPAVHSVTLSEBAAGAYAPLTAGTILINRVL 360
DB 301 PRALFASVRPGQVYVVAERDGRLLPAVHSVTLSEBAAGAYAPLTAGTILINRVL 360
QY 361 ASCYAIVEESHMAHRAFPRLAHLAALAPARTDRGSDGSGGGRGGRVALTPAGA 420
DB 361 ASCYAIVEESHMAHRAFPRLAHLAALAPARTDRGSDGSGGGRGGRVALTPAGA 420
QY 421 ADAPGAGATGIGHWYSQLYQIGTWLDSFALHPLGMAVSSXSXRGAGGARREGA 475
DB 421 ADAPGAGATGIGHWYSQLYQIGTWLDSFALHPLGMAVSSXSXRGAGGARREGA 475

RESULT 7

US-08-748-591-4
; Sequence 4, Application US/08748591
; Patent No. 5739811
; GENERAL INFORMATION:
; APPLICANT: Epstein, Ervin
; APPLICANT: Hu, Zhilan
; APPLICANT: Bonifas, Jeanette
; TITLE OF INVENTION: Mutant Human Hedgehog Gene
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish and Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,591
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: 06510/067001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-748-591-4

Query Match 97.5%; Score 2407; DB 1; Length 462;
Best Local Similarity 100.0%; Pred. No. 8.2e-233;

Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTARCLLVVSSLLVCSGLACGPGRGFGKRRHPKRLTPLAYKQFLPNVAEKTGLASG 60
DB 1 MLTARCLLVVSSLLVCSGLACGPGRGFGKRRHPKRLTPLAYKQFLPNVAEKTGLASG 60
QY 61 RYEEKISNSRERKELTPNYPDIIIFKDEENTGADRLMTORCKKLNALASVNNQWPGV 120
DB 61 RYEEKISNSRERKELTPNYPDIIIFKDEENTGADRLMTORCKKLNALASVNNQWPGV 120

DB 61 RYEEKISNSRERKELTPNYPDIIIFKDEENTGADRLMTORCKKLNALASVNNQWPGV 120
QY 121 KLRTVEGMDDEGHSEESLHYEGRAVDITTSDRDRSKYGMILARLAVBAGDWMYVESKAH 180
DB 121 KLRTVEGMDDEGHSEESLHYEGRAVDITTSDRDRSKYGMILARLAVBAGDWMYVESKAH 180
QY 181 IHCYKAKENSVAAKSGCCFGPSATVHLEOGCTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240
DB 181 IHCYKAKENSVAAKSGCCFGPSATVHLEOGCTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240
QY 241 FLDDDDAKKVFYIETREPERERLLTFAHLLFPAPHNDSATGEPKSSGSPSGALG 300
DB 241 FLDDDDAKKVFYIETREPERERLLTFAHLLFPAPHNDSATGEPKSSGSPSGALG 300
QY 301 PRALFASVRPGQVYVVAERDGRLLPAVHSVTLSEBAAGAYAPLTAGTILINRVL 360
DB 301 PRALFASVRPGQVYVVAERDGRLLPAVHSVTLSEBAAGAYAPLTAGTILINRVL 360
QY 361 ASCYAIVEESHMAHRAFPRLAHLAALAPARTDRGSDGSGGGRGGRVALTPAGA 420
DB 361 ASCYAIVEESHMAHRAFPRLAHLAALAPARTDRGSDGSGGGRGGRVALTPAGA 420
QY 421 ADAPGAGATGIGHWYSQLYQIGTWLDSFALHPLGMAVSS 462
DB 421 ADAPGAGATGIGHWYSQLYQIGTWLDSFALHPLGMAVSS 462

RESULT 8

US-08-748-591-9
; Sequence 9, Application US/08748591
; Patent No. 5739811
; GENERAL INFORMATION:
; APPLICANT: Epstein, Ervin
; APPLICANT: Hu, Zhilan
; APPLICANT: Bonifas, Jeanette
; TITLE OF INVENTION: Mutant Human Hedgehog Gene
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish and Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,591
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: 06510/067001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-748-591-9

Query Match 97.5%; Score 2407; DB 1; Length 462;
Best Local Similarity 100.0%; Pred. No. 8.2e-233;

Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MILLARCLLVVSSLLVCSGLACGPGRGFRKRRHPKKTPLAYKQFIPNVAEKTIGASG 60
DB 1 MILLARCLLVVSSLLVCSGLACGPGRGFRKRRHPKKTPLAYKQFIPNVAEKTIGASG 60
OY 61 REEGKISRSEFKELTPYVNPDIIFKDEENTGADRLMTORCKDKLNALATISVNMOMPGV 120
DB 61 REEGKISRSEFKELTPYVNPDIIFKDEENTGADRLMTORCKDKLNALATISVNMOMPGV 120
OY 121 KLRVTEGMDDEDHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYTESKAH 180
DB 121 KLRVTEGMDDEDHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYTESKAH 180
OY 181 IHCSYKAENSVAKSGGCPGSAVTHLEQGGTKLVKDLSPGDRVLAADDQGRILYSDFLT 240
DB 181 IHCSYKAENSVAKSGGCPGSAVTHLEQGGTKLVKDLSPGDRVLAADDQGRILYSDFLT 240
OY 241 FLDRDDGAKKVVYIETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSPSGGALG 300
DB 241 FLDRDDGAKKVVYIETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSPSGGALG 300
OY 301 PALFASRVPRGQRYVVAERDGRRLPAAVHSVTLSEEAAGAVAPLTAOGTILINRYL 360
DB 301 PALFASRVPRGQRYVVAERDGRRLPAAVHSVTLSEEAAGAVAPLTAOGTILINRYL 360
OY 361 ASCYAVIEHSHVAHRAFAFRLAHLAALAPARTDRGDSGGGGRGGRVALTAPGA 420
DB 361 ASCYAVIEHSHVAHRAFAFRLAHLAALAPARTDRGDSGGGGRGGRVALTAPGA 420
OY 421 ADAPGAGATAGIHWSQLLYOIGTWLDSSEALHPGLMAVKSS 462
DB 421 ADAPGAGATAGIHWSQLLYOIGTWLDSSEALHPGLMAVKSS 462

```

```

; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; TOPOLOGY: 1linear
; MOLECULE TYPE: protein
US-08-946-329A-20

Query Match      83.8%; Score 2068; DB 3; Length 437;
Best Local Similarity 87.4%; Pred. No. 7.5e-199;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

OY 1 MILLARCLLVVSSLLVCSGLACGPGRGFRKRRHPKKTPLAYKQFIPNVAEKTIGASG 60
DB 2 MILLARCLLVVSSLLVCSGLACGPGRGFRKRRHPKKTPLAYKQFIPNVAEKTIGASG 61
OY 61 REEGKISRSEFKELTPYVNPDIIFKDEENTGADRLMTORCKDKLNALATISVNMOMPGV 120
DB 62 REEGKISRSEFKELTPYVNPDIIFKDEENTGADRLMTORCKDKLNALATISVNMOMPGV 121
OY 121 KLRVTEGMDDEDHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYTESKAH 180
DB 122 KLRVTEGMDDEDHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYTESKAH 181
OY 181 IHCSYKAENSVAKSGGCPGSAVTHLEQGGTKLVKDLSPGDRVLAADDQGRILYSDFLT 240
DB 182 IHCSYKAENSVAKSGGCPGSAVTHLEQGGTKLVKDLSPGDRVLAADDQGRILYSDFLT 241
OY 241 FLDRDDGAKKVVYIETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSPSGGALG 300
DB 242 FLDRDDGAKKVVYIETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSPSGGALG 300
OY 301 PALFASRVPRGQRYVVAERDGRRLPAAVHSVTLSEEAAGAVAPLTAOGTILINRYL 360
DB 287 PALFASRVPRGQRYVVAERDGRRLPAAVHSVTLSEEAAGAVAPLTAOGTILINRYL 346
OY 361 ASCYAVIEHSHVAHRAFAFRLAHLAALAPARTDRGDSGGGGRGGRVALTAPGA 420
DB 347 ASCYAVIEHSHVAHRAFAFRLAHLAALAPARTDRGDSGGGGRGGRVALTAPGA 395
OY 421 ADAPGAGATAGIHWSQLLYOIGTWLDSSEALHPGLMAVKSS 462
DB 396 TEARGAEPFAGIHWSQLLYHIGTWLDSSEALHPGLMAVKSS 437

RESULT 10
US-08-567-357A-20
; Sequence 20, Application US/08567357A
; Patent No. 6132728
; GENERAL INFORMATION:
; APPLICANT: Beachy, Phillip A.
; APPLICANT: Moon, Randall T.
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; PRIOR APPLICATION DATA:
; FILING DATE: 07-OCT-1997
; APPLICATION NUMBER: 60/061,323
; FILING DATE: 07-OCT-1996
; APPLICATION NUMBER: 08/729,743
; FILING DATE: 10-JUL-1996
; APPLICATION NUMBER: 08/567,357
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/349,498
; FILING DATE: 02-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/140001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099

```

ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/080001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-567-357A-20

Query Match 83.8%; Score 2068; DB 4; Length 437;
Best Local Similarity 87.4%; Pred. No. 7, 5e-199;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLVVSSLVSGLAGPGRGFGRRHPPKLLPLAYKQIPNVAEKTLAGSG 60
DB 2 LLLARCFVLVTLASSLLVCPGLACGPGRGFGRRHPPKLLPLAYKQIPNVAEKTLAGSG 61
QY 61 RYEGKISRNSRFKELIPNYPNDIIFKDEENTGADRLMTQCKKLNLAISVNMOPGV 120
DB 62 RYEGKISRNSRFKELIPNYPNDIIFKDEENTGADRLMTQCKKLNLAISVNMOPGV 121
QY 121 KLRTGEGWDEGHSESLAHYEGRAVDITTSDRRSKYGMRLAVEAGFDWYYESKAH 180
DB 122 KLRTGEGWDEGHSESLAHYEGRAVDITTSDRRSKYGMRLAVEAGFDWYYESKAH 181
QY 181 IHCSVKAENSVAAKSGCGFPGSATVHLEQGCTKLVKDLSPDRVLAADDGRLLYSDFLT 240
DB 182 IHCSVKAENSVAAKSGCGFPGSATVHLEQGCTKLVKDLSPDRVLAADDGRLLYSDFLT 241
QY 241 FLDDDDGAKKFFVYIERPERERLLTAHLLFYAPHNDSATGPEASSGSGPPSGALG 300
DB 242 FLDDDDGAKKFFVYIERPERERLLTAHLLFYAPHNDSATGPEASSGSGPPSGALG 300
QY 301 PRALFASRVPRGQRYVVAERDGRLLPAVHSVTLSEEAAGAYAPLTAAGTILINRYL 360
DB 287 PSALFASRVPRGQRYVVAERDGRLLPAVHSVTLSEEAAGAYAPLTAAGTILINRYL 346
QY 361 ASCYAVIEEHSWAHRAFPFLAHLAALAPARTDRGDSGGDRGGGGRVALTPAGA 420
DB 347 ASCYAVIEEHSWAHRAFPFLAHLAALAPARTDRGDSGGDRGGGGRVALTPAGA 395
QY 421 ADAPGAGTAGIHWSQLLYIGTWLDSALHPGLMAVKS 462
DB 396 TEARGAEPTAGIHWSQLLYIGTWLDSALHPGLMAVKS 437

RESULT 11
US-08-729-743A-20
Sequence 20, Application US/08729743A
Patent No. 6214794
GENERAL INFORMATION:
APPLICANT: Beachy, Philip A.
APPLICANT: Moon, Randall T.
APPLICANT: Porter, Jeffrey A.
TITLE OF INVENTION: NOVEL HEDGEGHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95

SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,743A
FILING DATE: 07-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/567,357
FILING DATE: 04-DEC-1995
APPLICATION NUMBER: 08/349,498
FILING DATE: 02-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/099001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-729-743A-20

Query Match 83.8%; Score 2068; DB 4; Length 437;
Best Local Similarity 87.4%; Pred. No. 7, 5e-199;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLVVSSLVSGLAGPGRGFGRRHPPKLLPLAYKQIPNVAEKTLAGSG 60
DB 2 LLLARCFVLVTLASSLLVCPGLACGPGRGFGRRHPPKLLPLAYKQIPNVAEKTLAGSG 61
QY 61 RYEGKISRNSRFKELIPNYPNDIIFKDEENTGADRLMTQCKKLNLAISVNMOPGV 120
DB 62 RYEGKISRNSRFKELIPNYPNDIIFKDEENTGADRLMTQCKKLNLAISVNMOPGV 121
QY 121 KLRTGEGWDEGHSESLAHYEGRAVDITTSDRRSKYGMRLAVEAGFDWYYESKAH 180
DB 122 KLRTGEGWDEGHSESLAHYEGRAVDITTSDRRSKYGMRLAVEAGFDWYYESKAH 181
QY 181 IHCSVKAENSVAAKSGCGFPGSATVHLEQGCTKLVKDLSPDRVLAADDGRLLYSDFLT 240
DB 182 IHCSVKAENSVAAKSGCGFPGSATVHLEQGCTKLVKDLSPDRVLAADDGRLLYSDFLT 241
QY 241 FLDDDDGAKKFFVYIERPERERLLTAHLLFYAPHNDSATGPEASSGSGPPSGALG 300
DB 242 FLDDDDGAKKFFVYIERPERERLLTAHLLFYAPHNDSATGPEASSGSGPPSGALG 300
QY 301 PRALFASRVPRGQRYVVAERDGRLLPAVHSVTLSEEAAGAYAPLTAAGTILINRYL 360
DB 287 PSALFASRVPRGQRYVVAERDGRLLPAVHSVTLSEEAAGAYAPLTAAGTILINRYL 346
QY 361 ASCYAVIEEHSWAHRAFPFLAHLAALAPARTDRGDSGGDRGGGGRVALTPAGA 420
DB 347 ASCYAVIEEHSWAHRAFPFLAHLAALAPARTDRGDSGGDRGGGGRVALTPAGA 395
QY 421 ADAPGAGTAGIHWSQLLYIGTWLDSALHPGLMAVKS 462
DB 396 TEARGAEPTAGIHWSQLLYIGTWLDSALHPGLMAVKS 437

RESULT 12
US-09-057-860A-6
Sequence 6, Application US/09057860A
Patent No. 6277820
GENERAL INFORMATION:
APPLICANT: Arnon Rosenthal
APPLICANT: Mary Hynes
APPLICANT: Weilian Ye
TITLE OF INVENTION: Method Of Dopaminergic And Serotonergic
TITLE OF INVENTION: Neuron Formation From Neuroprogenitor Cells
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:

```

ADDRESSSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 Inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,860A
FILING DATE: 09-Apr-1998
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Syoboda, Craig G.
REGISTRATION NUMBER: 39,044
REFERENCE/DOCKET NUMBER: P1364
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1489
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-057-860A-6

```

```

Query Match      83.8%; Score 2068; DB 4; Length 437;
Best Local Similarity 87.4%; Pred. No. 7.5e-199;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

```

```

QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFGRRHPKLTPLAYKOFIPNVAEKTIGASG 60
DB 2 LLLARCFVLIVASSLLVPGIACGGRGFGRRHPKLTPLAYKOFIPNVAEKTIGASG 61
QY 61 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPGV 120
DB 62 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPGV 121
QY 121 KLRVTEGMEDDGHSEESLHYEGRAVDITTSRDRSKYGMLRLAVEGFDWVYTESKAH 180
DB 122 KLRVTEGMEDDGHSEESLHYEGRAVDITTSRDRSKYGMLRLAVEGFDWVYTESKAH 181
QY 181 IHCYKAENSVAAGSGCGPGSATVHLEOGGTRKLVKDSLPGDRLVAADOGRLYSDFLT 240
DB 182 IHCYKAENSVAAGSGCGPGSATVHLEOGGTRKLVKDLRPGDRLVAADOGRLYSDFLT 241
QY 241 FLDRDDGAKKVFYVETREPRERLLITAAHLLFVA PHND SATGEPPASSGSGPPSGGALG 300
DB 242 FLDRDDGAKKVFYVETREPRERLLITAAHLLFVA PHND SATGEPPASSGSGPPSGGALG 300
QY 301 PALFASRVRRPQORYVVAERGDRLRLPAVAHSTLSREAGATAPLTAOCTILINRLV 360
DB 302 PALFASRVRRPQORYVVAERGDRLRLPAVAHSTLSREAGATAPLTAOCTILINRLV 360
QY 361 ASCYAVIEESHVAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGGRVATLAPGA 420
DB 362 ASCYAVIEESHVAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGGRVATLAPGA 420
QY 421 ADAPGAGATAGIHWYSQLLYIGTWLDSSEALHPLGMAYKSS 462
DB 422 ADAPGAGATAGIHWYSQLLYIGTWLDSSEALHPLGMAYKSS 462
QY 462 ADAPGAGATAGIHWYSQLLYIGTWLDSSEALHPLGMAYKSS 462
DB 463 ADAPGAGATAGIHWYSQLLYIGTWLDSSEALHPLGMAYKSS 462

```

```

RESULT 13
US-08-349-498-20
Sequence 20, Application US/08349498
Patent No. 6281332
GENERAL INFORMATION:
APPLICANT: Beachy, Phillip A.
APPLICANT: Moon, Randall T.

```

```

TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: U.S.A.
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/349,498
FILING DATE: 02-DEC-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/043001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
US-08-349-498-20

```

```

Query Match      83.8%; Score 2068; DB 4; Length 437;
Best Local Similarity 87.4%; Pred. No. 7.5e-199;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

```

```

QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFGRRHPKLTPLAYKOFIPNVAEKTIGASG 60
DB 2 LLLARCFVLIVASSLLVPGIACGGRGFGRRHPKLTPLAYKOFIPNVAEKTIGASG 61
QY 61 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPGV 120
DB 62 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPGV 121
QY 121 KLRVTEGMEDDGHSEESLHYEGRAVDITTSRDRSKYGMLRLAVEGFDWVYTESKAH 180
DB 122 KLRVTEGMEDDGHSEESLHYEGRAVDITTSRDRSKYGMLRLAVEGFDWVYTESKAH 181
QY 181 IHCYKAENSVAAGSGCGPGSATVHLEOGGTRKLVKDSLPGDRLVAADOGRLYSDFLT 240
DB 182 IHCYKAENSVAAGSGCGPGSATVHLEOGGTRKLVKDLRPGDRLVAADOGRLYSDFLT 241
QY 241 FLDRDDGAKKVFYVETREPRERLLITAAHLLFVA PHND SATGEPPASSGSGPPSGGALG 300
DB 242 FLDRDDGAKKVFYVETREPRERLLITAAHLLFVA PHND SATGEPPASSGSGPPSGGALG 300
QY 301 PALFASRVRRPQORYVVAERGDRLRLPAVAHSTLSREAGATAPLTAOCTILINRLV 360
DB 302 PALFASRVRRPQORYVVAERGDRLRLPAVAHSTLSREAGATAPLTAOCTILINRLV 360
QY 361 ASCYAVIEESHVAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGGRVATLAPGA 420
DB 362 ASCYAVIEESHVAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGGRVATLAPGA 420
QY 421 ADAPGAGATAGIHWYSQLLYIGTWLDSSEALHPLGMAYKSS 462
DB 422 ADAPGAGATAGIHWYSQLLYIGTWLDSSEALHPLGMAYKSS 462
QY 462 ADAPGAGATAGIHWYSQLLYIGTWLDSSEALHPLGMAYKSS 462
DB 463 ADAPGAGATAGIHWYSQLLYIGTWLDSSEALHPLGMAYKSS 462

```

```

RESULT 14
PCT-US95-15463-20
Sequence 20, Application PC/TUS9515463

```



```

; GENERAL INFORMATION:
; APPLICANT: The Johns Hopkins University School of Medicine
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15463
; FILING DATE: 01-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/080M01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-15463-20

Query Match      83.8%; Score 2068; DB 5; Length 437;
Best Local Similarity 87.4%; Pred. No. 7.5e-199;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLVVSLVCSGLACGPGRGFKRRPKLTPLAYKQFIPIVAEKTIGASG 60
DB 2 LLLARCLVLVLASSLVCPGLACGPGRGFKRRPKLTPLAYKQFIPIVAEKTIGASG 61
QY 61 RYEGKISNSRFKELTPNYNPDIIFKDENTGADRLTQCKDKLNLAISVNMOPGV 120
DB 62 RYEGKITNSRFKELTPNYNPDIIFKDENTGADRLTQCKDKLNLAISVNMOPGV 121
QY 121 KLRVTEGDEGHHSEESLHTEGRAVDITTSDRDSKYGMLARLAVEGFPWVYVESKAH 180
DB 122 KLRVTEGDEGHHSEESLHTEGRAVDITTSDRDSKYGMLARLAVEGFPWVYVESKAH 181
QY 181 IHCSTKAENSYAAKSGCGFPGSATVNHLEOGGKTLVKDLPDGRVLAADDOGRLLYSDLT 240
DB 182 IHCSTKAENSYAAKSGCGFPGSATVNHLEOGGKTLVKDLPDGRVLAADDOGRLLYSDLT 241
QY 241 FLDRDGAKKVFYVETETEPREBRLTLTAHLLFVAPHNDSATGPEEASGSGPGSGALG 300
DB 242 FLDRDGAKKVFYVETETEPREBRLTLTAHLLFVAPHNDSATGPEEASGSGPGSGALG 300
QY 301 PRALEFASVVRGQRYVVAERDGRRLPAAVHSTVLTSEEAAGAPLTAOCTILINRYL 360
DB 287 PSALFASVVRGQRYVVAERDGRRLPAAVHSTVLTSEEAAGAPLTAOCTILINRYL 366
QY 361 ASCYAVIEHSHMAHRAFPRLAHALLAALAPARTDRGDSGGGCGGGRVALTARGA 420
DB 347 ASCYAVIEHSHMAHRAFPRLAHALLAALAPARTDRGDSGGGCGGGRVALTARGA 420
QY 421 ADAPGATAGIHWSQLLYOIGTWLDSSEALHPGLMAVKSS 462
DB 396 TEARGAEPTAGIHWSQLLYOIGTWLDSSEALHPGLMAVKSS 437

RESULT 15
```

```

PCT-US95-15923-20
; Sequence 20, Application PC/US9515923
; GENERAL INFORMATION:
; APPLICANT: The Johns Hopkins University School of Medicine, et al.
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15923
; FILING DATE: 04-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/043M01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-15923-20

Query Match      83.8%; Score 2068; DB 5; Length 437;
Best Local Similarity 87.4%; Pred. No. 7.5e-199;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLVVSLVCSGLACGPGRGFKRRPKLTPLAYKQFIPIVAEKTIGASG 60
DB 2 LLLARCLVLVLASSLVCPGLACGPGRGFKRRPKLTPLAYKQFIPIVAEKTIGASG 61
QY 61 RYEGKISNSRFKELTPNYNPDIIFKDENTGADRLTQCKDKLNLAISVNMOPGV 120
DB 62 RYEGKITNSRFKELTPNYNPDIIFKDENTGADRLTQCKDKLNLAISVNMOPGV 121
QY 121 KLRVTEGDEGHHSEESLHTEGRAVDITTSDRDSKYGMLARLAVEGFPWVYVESKAH 180
DB 122 KLRVTEGDEGHHSEESLHTEGRAVDITTSDRDSKYGMLARLAVEGFPWVYVESKAH 181
QY 181 IHCSTKAENSYAAKSGCGFPGSATVNHLEOGGKTLVKDLPDGRVLAADDOGRLLYSDLT 240
DB 182 IHCSTKAENSYAAKSGCGFPGSATVNHLEOGGKTLVKDLPDGRVLAADDOGRLLYSDLT 241
QY 241 FLDRDGAKKVFYVETETEPREBRLTLTAHLLFVAPHNDSATGPEEASGSGPGSGALG 300
DB 242 FLDRDGAKKVFYVETETEPREBRLTLTAHLLFVAPHNDSATGPEEASGSGPGSGALG 300
QY 301 PRALEFASVVRGQRYVVAERDGRRLPAAVHSTVLTSEEAAGAPLTAOCTILINRYL 360
DB 287 PSALFASVVRGQRYVVAERDGRRLPAAVHSTVLTSEEAAGAPLTAOCTILINRYL 366
QY 361 ASCYAVIEHSHMAHRAFPRLAHALLAALAPARTDRGDSGGGCGGGRVALTARGA 420
DB 347 ASCYAVIEHSHMAHRAFPRLAHALLAALAPARTDRGDSGGGCGGGRVALTARGA 420
QY 421 ADAPGATAGIHWSQLLYOIGTWLDSSEALHPGLMAVKSS 462
DB 396 TEARGAEPTAGIHWSQLLYOIGTWLDSSEALHPGLMAVKSS 437
```

Thu Feb 20 10:43:53 2003

us-09-827-110a-15.ra1

Page 10

Search completed: February 20, 2003, 10:14:13
Job time : 14.4346 secs

...

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:08:18 ; Search time 15.632 Seconds
(without alignments)
2921.176 Million cell updates/sec

Title: US-09-827-110A-15
Perfect score: 2469
Sequence: 1 MLILARCLLVLVSSLLVCS.....GMAVSSKSRGAGCGAREGA 475

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2062	83.5	437	2 A49425	Sonic hedgehog pro
2	2060	83.4	437	2 B53193	hedgehog homolog v
3	1807	73.2	425	2 A49424	patterning protein
4	1506	61.0	418	2 A53193	hedgehog homolog v
5	1498	60.7	444	2 S56765	morphogen Xhh prec
6	1459	59.1	415	2 A49426	sonic hedgehog gen
7	1183	47.9	396	2 B49425	Desert hedgehog pr
8	1027.5	41.6	336	2 C49425	Indian hedgehog pr
9	890	36.0	471	2 A46400	segment polarity p
10	449	18.2	94	2 G02735	desert hedgehog -
11	159.5	6.5	615	2 T29550	hypothetical prote
12	149	6.0	1207	2 T23754	hypothetical prote
13	138.5	5.6	1226	2 T24045	hypothetical prote
14	135.5	5.5	484	2 T34504	hypothetical prote
15	134.5	5.4	481	2 T27665	hypothetical prote
16	129	5.2	1147	1 MMAX1B	myosin heavy chain
17	128	5.2	629	2 T19563	hypothetical prote
18	126.5	5.1	1021	2 T23252	hypothetical prote
19	123.5	5.0	2639	2 T31328	fibronin - Chinese
20	121	4.9	868	2 T22281	hypothetical prote
21	119	4.8	318	2 T35213	probable secreted
22	118.5	4.8	702	2 G01840	T-box protein 2 -
23	116	4.7	1329	2 E70917	hypothetical glyci
24	115.5	4.7	940	2 E87250	(protein-PII) urid
25	115	4.7	486	2 H87311	hypothetical glyci
26	114	4.6	741	2 G70917	hypothetical prote
27	111	4.5	333	2 A39065	homeotic protein E
28	111	4.5	1298	1 EDBE75	Immediate-early pr
29	110	4.5	462	2 T35852	Probable UDP-N-ace

30	107.5	4.4	701	1 S46458	transcription fact
31	107	4.3	775	1 EDBE11	Immediate-early pr
32	106.5	4.3	482	2 E70554	hypothetical prote
33	106	4.3	482	2 JC7583	basic helix-loop-h
34	105.5	4.3	490	2 J09084	phosphatidylinosit
35	105.5	4.3	841	2 C87331	ISCC2, transposase
36	105.5	4.3	1324	2 T14070	peptide synthetase
37	105.5	4.3	7463	2 T36248	CDA peptide synthe
38	105	4.3	460	2 T33110	hypothetical prote
39	104.5	4.2	510	2 A55207	glycerol-3-phospha
40	104	4.2	402	1 S25077	monensin polyketid
41	104	4.2	1063	2 A40253	acidic nuclear pro
42	104	4.2	3161	2 T30342	protein HMW1 - Ye
43	103.5	4.2	975	2 S33121	homeotic protein C
44	102.5	4.2	413	2 E70661	probable PE protei
45	102.5	4.2	479	2 F70573	hypothetical prote

ALIGNMENTS

RESULT 1

A49425
Sonic hedgehog protein precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999

C:Accession: A49425

R:Reheland, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; M

A:Title: Sonic hedgehog, a member of a family of putative signalling molecules, is imp

A:Reference number: A49425; M0ID:94094334; PMID:7916661

A:Accession: A49425

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1437 <ECH>

A:Cross-references: GB:X76290

A>Note: authors translation is shown for the codon TCC at position 436

C:Genetics:

A:Gene: Shh

C:Superfamily: sonic hedgehog protein

Query Match 83.5%; Score 2062; DB 2; Length 437;
Best Local Similarity 87.0%; Pred. No. 3.6e-147;
Matches 402; Conservative 10; Mismatches 24; Indels 26; Gaps 3;

QY	1	MLILARCLLVIVSLVCSGLACGPGRGKRRPKPLPLAYKOTIPVAEKTGASG	60
DB	2	LLILARCLLVIVSLVCSGLACGPGRGKRRPKPLPLAYKOTIPVAEKTGASG	61
QY	61	RYEKGISRNSERFKELTPNTNPDIIFKDEBNTGADRLMTORCKDLNALAISVNMOPGV	120
DB	62	RYEKGITRNSERFKELTPNTNPDIIFKDEBNTGADRLMTORCKDLNALAISVNMOPGV	121
QY	121	KLRVTEGWDDEGHHSESLYEGRAVDITTSDRDSRYGMLARLAVAGCDWYESSKAH	180
DB	122	KLRVTEGWDDEGHHSESLYEGRAVDITTSDRDSRYGMLARLAVAGCDWYESSKAH	181
QY	181	IHCYKAKENSVAAKSGCGPGSATVHLEGGTKYKRLSPGDRVLADDDGRLLYSFLT	240
DB	182	IHCYKAKENSVAAKSGCGPGSATVHLEGGTKYKRLSPGDRVLADDDGRLLYSFLT	241
QY	241	FLRDDGAKKVFYVIEFREPRLLLTAHLLEFVAPHNDSATGEPKASGSGPPSGCALG	300
DB	242	FLRDDGAKKVFYVIEFREPRLLLTAHLLEFVAPHNDSATGEPKASGSGPPSGCALG	301
QY	301	PRALFASRVPRGQRYVVAERDGRLLPAVHSVTLSSEAAGYADLTAGCTLLIRVL	360
DB	287	PSALFASRVPRGQRYVVAERDGRLLPAVHSVTLSSEAAGYADLTAGCTLLIRVL	346
QY	361	ASCAVVEEHSMAHRAFPRLAHLAALAPATDGGSGGDRGGGGGRVLTAPGA	420
DB	347	ASCAVVEEHSMAHRAFPRLAHLAALAPATDGGSGGDRGGGGGRVLTAPGA	395

A:Cross-references: GB:X76292; NID:9443941; PID:CAA53924.1; PID:9443942
 C:Genetics:
 A:Gene: Dhh
 C:Superfamily: sonic hedgehog protein

Query Match 47.9%; Score 1183; DB 2; Length 396;
 Best Local Similarity 54.5%; Pred. No. 3.3e-81;
 Matches 244; Conservative 55; Mismatches 87; Indels 62; Gaps 10;

```

OY 3 LLARCLLVYSSLLVCSGLACGPGRG-FGKRHRPK-LTPAYKQFIPNVAEKLIGASG 60
DB 7 LILPLCL-----ALALASGSGPGRGVGRRRYRKQVLPLTYQFYSMERLIGASG 61
OY 61 RYEGATSRNRSERFKELTPNYNDIIFKDEBNTGADRLMTQRCCKDLALAISSVNMOWPGV 120
DB 62 PAEGVTVTSGSERFDLVNPNYNDIIFKDEBNSGADRLMTQRCCKERYNMAIAIYMMWMPGV 121
OY 121 KLRTVEGDEQHHSEESIHYECRAVDITTSDRDRSKYGMRLAVEGDFWVYESKAH 180
DB 122 RLRTVEGDEQHHADSDLIHGRALDITTSDRDRNKYGLRLAVEGDFWVYESRHH 181
OY 181 IHCYKAENSVAAKSGGPGSATVHLEQGGTKLVKDISPCGRVLAADQGRILYSDFLT 240
DB 182 IHVSFKADNSLAVRAGCGEPGNATVRLSRGKGLRELHKGWVLAADAAGNVPTVILL 241
OY 241 FLDRDDGAKKVFYVIEETREPRRLITAAHLLEFVAHPHDSATGEPAASSGSGP-PSGAL 299
DB 242 FLDRDLQRRASFVAVETERPRLKLLTPMHLVFAA-----RGPAPAPGDF 286
OY 300 GPRAFFASRRVRCGRVYVAERDGRRLPPAAVHVTLSSEPAAGVAPLTAGOTLLINRV 359
DB 287 AP--VFARRLRAGDSVLA---PGSDALQPARAVARA-REEVAVGAPLTAGTLLVNDV 339
OY 360 LASCYAVIEHSHMARAFAPFRLAHLAALPARTDRGDSGGDGGGGGVALTAPG 419
DB 340 LASCYAVIESHQMAHARAPRLAHLALGALL-----PG 372
OY 420 AADAFAGATACIHWYSOLLVOIGTWLL 447
DB 373 GAVOP-----TGMWYSRLLYLAEEELM 395

```

RESULT 8

A:Cross-references: GB:X76292; NID:9443941; PID:CAA53924.1; PID:9443942
 C:Genetics:
 A:Gene: Dhh
 C:Superfamily: sonic hedgehog protein

Query Match 41.6%; Score 1027.5; DB 2; Length 336;
 Best Local Similarity 54.6%; Pred. No. 1.3e-69;
 Matches 218; Conservative 39; Mismatches 77; Indels 65; Gaps 8;

```

OY 71 ERFKELPNYPDIIFKDEBNTGADRLMTQRCCKDLALAISSVNMOWPGVLAETEGWDE 130
DB 1 ERFKELPNYPDIIFKDEBNTGADRLMTQRCCKDLALAISSVNMOWPGVLAETEGWDE 60
OY 131 DGHSESESLAYEGRAVDITTSDRDRSKYGMRLAVEGDFWVYESKAHICSVKANS 190
DB 61 DGHSESESLAYEGRAVDITTSDRDRNKYGLRLAVEGDFWVYESKAHVCSVSEHS 120
OY 191 VAAKSGGPGSATVHLEQGGTKLVKDISPCGRVLAADQGRILYSDFLTFLDRDDGAKK 250

```

```

DB 121 AAKTGGCGPAGAOVLENGERVALSAVPGDRIAMGEDGTPFSDVLIFLDRPNRLR 180
OY 251 VFYIETREPRERLITAAHLLEFVAHPHDSATGEPAASSGSGPFGSGALGPALPASRV 310
DB 181 AFQVETQDPPRRALITRAHLLEFIA---DNHT-EPAAHF-----RATASHVO 224
OY 311 PGORVYVAERDGRRLPPAAVHVTLSSEPAAGVAPLTAGOTLLINRVLASCVVIEBH 370
DB 225 PQQVYLV-----SGVGLQPARVAAVS-THVALGSYAPLTRHGTLVEDVAVSCFAVVDH 279
OY 371 SWAHAFAPFRLAHL--AALAPARTDRGDSGGDGGGGGVALTAPGADAGACAT 429
DB 280 HLAQLAFWPLRLEPFLANGSWTPSE----- 304
OY 430 AGIHWYSOLLVOIGTWLLDSEALHPLGMAVKSXSRGAG 468
DB 305 -GVHSYPMQMLYRLGRLLLESTFHPLGMS-----GAG 335

```

RESULT 9

A:Cross-references: GB:X76292; NID:9443941; PID:CAA53924.1; PID:9443942
 C:Genetics:
 A:Gene: Dhh
 C:Superfamily: sonic hedgehog protein

Query Match 36.0%; Score 890; DB 2; Length 471;
 Best Local Similarity 44.3%; Pred. No. 4.3e-59;
 Matches 202; Conservative 65; Mismatches 137; Indels 52; Gaps 10;

```

OY 6 RCL-----LVVYSSLLVCSGLACGPGRGFGKRRRPKLTPLAYKQFIPNVAEKLIG 57
DB 59 RCLSRITSLVALITLIVLPVSPSPAHSCGPGRGRLGRRH-ARNILYPLVLYKQFIPNLSYTN 117

```

[illegible]

```

RESULT 10
G02735
desert hedgehog - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 11-Jan-2000
C:Accession: G02735
R:Drummond, I.A.
submitted to the EMBL Data Library, June 1996
A:Reference number: H01643
A:Accession: G02735
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-94 <DRU>
A:Cross-references: EMBL:U59748; NID:g1401271; PID:g1401272
C:Genetics:
A:Gene: hDHH
C:Superfamily: sonic hedgehog protein

Query Match      18.2%; Score 449; DB 2; Length 94;
Beat Local Similarity 84.0%; Pred. No. 7e-27;
Matches 79; Conservative 13; Mismatches 2; Indels 0; Gaps 0;

OY 84 IIFKDEEYTGADRLTQRCCKDKLNLAIISVNWQMGVKIRVTEGDEDSHSEESLHREG 143
      |||FKEENS|GADRLTECKE|KERNVNLAIAYNMMPGVRLRYTEGWEDG|HNAQDSLH|YEG 60
Db 1

OY 144 RAVDITTS|DRDRSKY|GMLARLAVEAGFPWVYVES 177
      |||:|||||:|||||:|||||:|||||:|||||:|
Db 61 RALDITTS|DRDRNKKY|GLARLAVEAGFDWYVYGS 94

RESULT 11
T29550
hypothetical protein ZK377.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T29550
R:Nhan, M.; Hawkins, J.
submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans coamid ZK377.
A:Reference number: Z20639
A:Accession: T29550
A:Status: preliminary; translated from GB/EMBL/DBJ

```

A:Molecule type: DNA
A:Residues: 1-615 <NHAS>
A:Cross-references: EMBL:U088183, PUDN:AA52656.1, GSPDB:GN00028, CESP:ZK377.1
A:Experimental source: strain Bristol N2; clone ZK377
C:Genetics:
A:Gene: CESP:ZK377.1
A:Map position: X
A:Introns: 46/1; 70/3; 157/2; 176/1; 245/2; 314/3; 402/3; 499/2

Query Match	6.5%;	Score 159.5;	DB 2;	Length 615;
Best Local Similarity	28.3%;	Pred. No. 0.00044;		
Matches	52;	Conservative	33;	Mismatches 74; Indels 25; Gaps 7;
QY	198	CFPGSAYHLEGGGKTLYKDS	PSGGRVLAADQGRLLYSDFELTFRDDGAAKVEYUET	257
Db	416	CFSGMEVEYETD-GIKMKTLKIGKRLVSM-D	DEAFVTYSPVIMFLHKRDEIAEENLLET	473
QY	258	REPRERLLTLTAHLLEFVAHPHDSATGCEPDEASSGSP	PSGALGRPALFASVRPCGRVYV	317
Db	474	ANGHS-IKLTIDNHLTY--SDCFTRS-----	DELVAKEKVKMDCCIV	514
QY	318	VAERDGRRLRLPAAYHVSYTLSEEAAGAYAPRLTAOCTIL	INRVYASCYUVEEHSMAHRAV	377
Db	515	TTD--SNWVTKKKVSKISKVTE-TGISPSPLSTGDIIVN	RYLASCNSYLAKSLQGRFF	570
QY	378	APFR 381		
Db	571	SLVK 574		

[illegible]

Db 1053 VAMRESKAEARKCEVLSIDE-SGE-----VIADEIVRGRTNVGIYSPMT 1100
QY 350 AOGTILINRVLASCYAVIEHSMARAPAF 380
Db 1101 VEGSLIVDGVLLSCFSHLSHS-AHKLFEDF 1130

RESULT 13

T24045

hypothetical protein R08B4.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T24045

R:White, S.
submitted to the EMBL Data Library, November 1995

A:Reference number: Z19834

A:Accession: T24045

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1226 <WIL>

A:Cross-references: EMBL:Z68008; PIDN:CAA92000.1; GSPDB:GN00028; CESP:R08B4.1

A:Experimental source: clone R08B4

C:Genetics:

A:Gene: CESP:R08B4.1

A:Map position: X

A:Introns: 53/2; 113/1; 152/3; 204/3; 272/1; 354/1; 389/3; 714/3; 839/3; 877/1; 960/3; 1

Query Match 5.6%; Score 138.5; DB 2; Length 1226;

Best Local Similarity 24.4%; Pred. No. 0.041;

Matches 50; Conservative 38; Mismatches 68; Indels 49; Gaps 8;

QY 189 NSVAAKSGCEPPGSAVHLEGGTKLVKLSPGDRVLAADOGRLYSDFLTDRDGA 248
Db 1012 SALVATGACFSIDTWV-TTFPGKRRMDIDIGDYVLTADLE-KYFPPTLMIHREPK 1069
QY 249 KKVIVITRPRERLTLTAHLF-----VAHNSATGEPPEASGSGSPS 295
Db 1070 VQEFFTIT-EGKTLRTSRFMYRNKCGKSYQYIKMLPD-----GE----- 1113QY 296 GCGALPRALFASRVPGQRYVVAERDGRLLPAVHVSVTLSEBAGAVAPLTAQGTIL 355
Db 1114 -----AIFSDLEVGCYVVLVYKGRKQKIEITRSV-----RTGISPLTNKRII 1161QY 356 INRVLASCYAVIEHS-----MAH 374
Db 1162 VNDMLASCYSEIQNTLQTFEMAY 1186

RESULT 14

T34504

hypothetical protein ZK1290.12 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T34504

R:Taich, A.
submitted to the EMBL Data Library, July 1995

A:Description: The sequence of C. elegans cosmid ZK1290.

A:Reference number: Z21535

A:Accession: T34504

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-484 <TAI>

A:Cross-references: EMBL:U21308; PIDN:AAB93321.1; GSPDB:GN00020; CESP:ZK1290.12

A:Experimental source: strain Bristol N2; clone ZK1290

C:Genetics:

A:Gene: CESP:ZK1290.12

A:Map position: 2

A:Introns: 66/3; 124/2; 181/2; 392/1

Query Match 5.5%; Score 135.5; DB 2; Length 484;

Best Local Similarity 24.5%; Pred. No. 0.021;

Matches 46; Conservative 42; Mismatches 65; Indels 35; Gaps 8;

QY 191 VAAKSGCEPPGSAVHLEGGTKLVKLSPGDRVLAADOGRLYSDFLTDRDGA 250
Db 283 VASGVPACTGNSKV-KTPAGEKSMADLSVGDMMVTV-EGKMTYTRVASMILRLPDRKA 340
QY 251 VFVYETREPRERLTLTAHLFVAPPHNSATGEPPEASGSGPPGALPRALFASRV 310
Db 341 AFIKL-TTEOGAIDMTPOHFIYKA---NCVTEME-----LVYAEEDMT 380
QY 311 PGQRYVVAERDGRLLPAVHVSVTLSEBAGAVAPLTAQGTILINRVLASCYAV 366
Db 381 IGDCLMV---KENELVM-----TTISEKSTFYETGYAPRPTETGDLIVDYVASCINAV 431
QY 367 IEHSHWAH 374
Db 432 VKANTLSH 439

RESULT 15

T27665

hypothetical protein ZK1037.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27665

R:Baisham, V.
submitted to the EMBL Data Library, October 1996

A:Reference number: Z20401

A:Accession: T27665

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-481 <WIL>

A:Cross-references: EMBL:Z81142; PIDN:CA803509.1; GSPDB:GN00023; CESP:ZK1037.10

A:Experimental source: clone ZK1037

C:Genetics:

A:Gene: CESP:ZK1037.10

A:Map position: 5

A:Introns: 37/3; 75/3; 141/3; 173/2; 253/1; 298/3; 310/1; 371/2; 424/3

Query Match 5.4%; Score 134.5; DB 2; Length 481;

Best Local Similarity 24.3%; Pred. No. 0.024;

Matches 45; Conservative 34; Mismatches 67; Indels 39; Gaps 6;

QY 198 CFPGSAVHLEGGTKLVKLSPGDRVLAADOGRLYSDFLTLD-----RDGAKKVF 252
Db 314 CFPNDVAVNVYEKAVKRMDELEIGDMVALDENG-----EDITFLPVKYLHROPEQAE 368QY 253 VVETREPRERLTLTAHLFVAPPHNSATGEPPEASGSGPPGALPRALFASRVPG 312
Db 369 FLERSLDNGETFTLEKHLVY-----TTECRONS-----SEKISWESISAGKVAVG 415QY 313 QRVYVVAERDGRLLPAVHVSVTLSE-----ENAGAVAPLTAQGTILINRVLASCYA 365
Db 416 DCFYLAQSE-----ALKRYRLVELLDIKRVKKTQIYAPMNSQGHLLVKNKITSCHS 466QY 366 VIEEH 370
Db 467 EVDHH 471

Search completed: February 20, 2003, 10:13:37
Job time : 18.632 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:07:32 ; Search time 14.5662 seconds
(without alignments)

1352.533 Million cell updates/sec

Title: US-09-827-110A-15

Perfect score: 2469
Sequence: 1 MLILARCLLLVLSILVCS.....GNAYKSKSRGAGGAREGA 475

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2407	97.5	462	1 SHH_HUMAN	015465 homo sapien
2	2068	83.8	437	1 SHH_MOUSE	062226 mus musculu
3	2060	83.4	437	1 SHH_RAT	063673 rattus norv
4	1807	73.2	425	1 SHH_CHICK	091075 gallus gall
5	1539.5	62.4	432	1 SHH_CYNPY	090385 cynops pyr
6	1506	61.0	418	1 SHH_BRARE	092008 brachydanto
7	1498	60.7	444	1 SHH_XENLA	092000 xenopus lae
8	1432	58.0	416	1 TWHH_BRARE	090419 brachydanto
9	1282.5	51.9	408	1 IHH_CHICK	098938 gallus gall
10	1266	51.3	411	1 IHH_HUMAN	014623 homo sapien
11	1262	51.1	411	1 IHH_MOUSE	097812 mus musculu
12	1248	50.5	409	1 IHH_XENLA	091612 xenopus lae
13	1213	49.1	412	1 IHH_BRARE	098862 brachydanto
14	1183	47.9	386	1 DHH_MOUSE	061488 mus musculu
15	1177	47.7	386	1 DHH_HUMAN	043333 homo sapien
16	1079.5	43.7	396	1 DHH1_XENLA	091610 xenopus lae
17	1077.5	43.6	398	1 DHH2_XENLA	091611 xenopus lae
18	890	36.0	471	1 HH_DROME	002936 drosophila
19	854	34.6	481	1 HH_DROHY	056674 drosophila
20	602	24.4	121	1 SHH_RASEL	079858 rasbora ele
21	599	24.3	121	1 SHH_CARAU	079691 carassius a
22	599	24.3	121	1 SHH_PUNTE	079850 punctus tet
23	598	24.2	121	1 SHH_RASHE	079864 rasbora het
24	598	24.2	121	1 SHH_RASPA	079869 rasbora pav
25	597	24.2	121	1 SHH_TANAL	079915 tanichthys
26	595	24.1	121	1 SHH_AMBCH	079682 amblypharyn
27	594	24.1	121	1 SHH_DANNA	013235 danio aff.
28	594	24.1	121	1 SHH_DANNA	013234 danio aff.
29	594	24.1	121	1 SHH_DANAT	013238 danio aff.
30	594	24.1	121	1 SHH_DANAT	013245 danio frank
31	594	24.1	121	1 SHH_DANKE	079709 danio kerri
32	594	24.1	121	1 SHH_DANKE	079717 danio pulch
33	594	24.1	121	1 SHH_DEVDE	013241 devario dev

34	594	24.1	121	1 SHH_DEVMA	013247 devario mal
35	594	24.1	121	1 SHH_DEVPA	013250 devario pat
36	590	23.9	121	1 SHH_PUNCO	079838 punctus con
37	396	16.0	88	1 DHH_BRARE	079729 brachydanto
38	293	11.9	58	1 IHH_CARAU	079693 carassius a
39	293	11.9	58	1 IHH_DANAT	013240 danio aff.
40	293	11.9	58	1 IHH_DANKE	079711 danio kerri
41	293	11.9	58	1 IHH_DANKE	079719 danio kerri
42	293	11.9	58	1 IHH_DEVDE	013243 devario dev
43	293	11.9	58	1 IHH_PUNTE	079852 punctus tet
44	293	11.9	58	1 IHH_RASEL	079860 rasbora ele
45	289	11.7	58	1 SHH_PSEPR	079839 pseudorasbo

ALIGNMENTS

RESULT 1
ID SHH_HUMAN STANDARD; PRT; 462 AA.
AC 015465;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sonic hedgehog protein precursor (SHH) (HMG-1).
GN SHH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=96070431; PubMed=7590746;
RA Marigo V., Roberts D.J., Lee S.M.K., Tsukurov O., Levi T.,
RA Gastier J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seldman C.E.,
RA Jenkins N.A., Seldman J.G., McMahon A.P., Tabin C.;
RT "Cloning, expression, and chromosomal location of SHH and IHH: two
RT human homologues of the Drosophila segment polarity gene hedgehog.";
RL Genomics 28:44-51(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Tate G., Kishimoto K., Mitsuura T.;
RT "Expression of Sonic hedgehog and its receptor Patched/Smoothed in
RT human cancer cell lines and embryonic organs.";
RL J. Biochem. Mol. Biol. Biophys. 4:27-34(2000).
RN [3]
RP SEQUENCE OF 1-187 FROM N.A.
RA Strong C., Graves T., Sutterer C., Ozersky P.;
RT Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 119-167 FROM N.A.
RX MEDLINE=95236997; PubMed=7720571;
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Stmandl B.K.,
RA Zhao R., Seidman M.F., Fallon J.F., Beachy P.A.;
RT "Products, genetic linkage and limb patterning activity of a murine
RT hedgehog gene.";
RL Development 120:3339-3353(1994).
RN [5]
RP PALMITOYLATION.
RX MEDLINE=98256339; PubMed=9593755;
RA Peplinsky R.B., Zeng C., Wen D., Rayhorn P., Baker D.P., Williams K.P.,
RA Bixler S.A., Ambrose C.M., Garber E.A., Matkowski K., Taylor F.R.,
RA Wang E.A., Gales A.;
RT "Identification of a palmitic acid-modified form of human Sonic
RT hedgehog.";
RL J. Biol. Chem. 273:14037-14045(1998).
RN [6]
RP VARIANTS HPE ARG-31; GLY-117 AND ARG-117.
RX MEDLINE=97051937; PubMed=8896572;
RA Roessler E., Belloni E., Gaudenz K., Jay P., Berta P., Scherer S.W.,
RA Tsui L.-C., Muenke M.;
RT "Mutations in the human Sonic hedgehog gene cause holoprosencephaly.";

RL Nat. Genet. 14:357-360(1996).
 RP [17]
 RP VARIANTS HPE ARG-31; GLY-117; ARG-117; GLU-224; THR-226 AND THR-383.
 RX MEDLINE-96027056; PubMed-9302262;
 RA Roessler E., Belloni E., Gaudenz K., Vargas F., Scherer S.W.,
 RA Tsai L.-C., Muenke M.,
 RT "Mutations in the C-terminal domain of Sonic hedgehog cause
 RT holoprosencephaly.";
 RL Hum. Mol. Genet. 6:1847-1853(1997).
 RN [8]
 RP VARIANTS HPE HIS-100; GLN-188 AND ASN-222.
 RX MEDLINE-99371775; PubMed-10441331;
 RA Odent S., Attli-Bitach T., Blayau M., Mathieu M., Aug J.,
 RA Delezo de A.V., Gall J.Y., Le Marec B., Munnich A., David V.,
 RA Vekemans M.,
 RT "Expression of the sonic hedgehog (SHH) gene during early human
 RT development and phenotypic expression of new mutations causing
 RT holoprosencephaly.";
 RL Hum. Mol. Genet. 8:1683-1689(1999).
 RN [9]
 RP VARIANTS HPE V-88; K-115; R-236; 263-R--A-269 DEL; D-290; A-424 AND
 RP L-436.
 RX MEDLINE-20025757; PubMed-10556296;
 RA Nanni L., Ming J.E., Bocian M., Steinhaus K., Bianchi D.M.,
 RA Die-Smulders C., Giannotti A., Imailumi K., Jones K.L., Campo M.D.,
 RA Martin R.A., Melnick P., Pierpont M.E.M., Robin N.H., Young I.D.,
 RA Roessler E., Muenke M.,
 RT "The mutational spectrum of the sonic hedgehog gene in
 RT holoprosencephaly: SHH mutations cause a significant proportion of
 RT autosomal dominant holoprosencephaly.";
 RL Hum. Mol. Genet. 8:2479-2488(1999).
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
 CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
 CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN PERAL INTESTINE, LIVER, LUNG, AND
 CC KIDNEY. NOT EXPRESSED IN ADULT TISSUES.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- DISEASE: DEFECTS IN SHH ARE THE CAUSE OF HOLOPROSENCEPHALY (HPE).
 CC HPE IS A CLINICALLY VARIABLE AND GENETICALLY HETEROGENEOUS
 CC MALFORMATION IN WHICH THE DEVELOPING FOREBRAIN FAILS TO CORRECTLY
 CC SEPARATE INTO RIGHT AND LEFT HEMISPHERES. IN THE MOST SEVERE FORM
 CC (ALOBAR HPE), THERE IS NO INTERHEMISPHERIC FISSURE, A SINGLE BRAIN
 CC VENTRICLE IS PRESENT. HPE IS ASSOCIATED WITH SEVERAL DISTINCT
 CC FACIES AND PHENOTYPIC VARIABILITY. IN THE MOST EXTREME CASES,
 CC ANOPHTHALMIA OR CYCLOPIA IS EVIDENT ALONG WITH A CONGENITAL
 CC ABSENCE OF THE MATURE NOSE. THE LESS SEVERE FORM FEATURES FACIAL
 CC DYSMORPHIA CHARACTERIZED BY OCULAR HYPERTELORISM, DEFECTS OF THE
 CC UPPER LIP AND/OR NOSE, AND ABSENCE OF THE OLFACTORY NERVES OR
 CC CORPUS CALLOSUM. THE MAJORITY OF HPE CASES ARE APPARENTLY
 CC SPORADIC, ALTHOUGH CLEAR EXAMPLES OF AUTOSOMAL DOMINANT (AD)
 CC INHERITANCE HAVE BEEN DESCRIBED. INTERESTINGLY, UP TO 30% OF

CC OBLIGATE CARRIERS OF HPE GENE IN AD PEDIGREES ARE CLINICALLY
 CC UNAFFECTED.
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -1- DATABASE: NAME-Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW="http://www.intoblogen.fr/services/chromosome/Genes/SHHID378.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 CC EMBL: L38518; AAA62179.1; -
 CC HSPB: AC002484; AAB67604.1; -
 CC HSSP: O62226; 1YH.
 CC DR MEMOPS; C46.002; -
 CC DR Genew; HGNC:10848; SHH.
 CC DR MIM; 600725; -
 CC DR MIM; 142945; -
 CC DR InterPro: IPR000320; HH_signal.
 CC DR InterPro: IPR001767; Hedgehog_hnt.
 CC DR InterPro: IPR003586; Hedgehog_hntC.
 CC DR InterPro: IPR003587; Hedgehog_hntN.
 CC DR InterPro: IPR002203; Intein.
 CC DR InterPro: IPR001657; SonicHH.
 CC DR Pfam: PF01079; Hnt; 1.
 CC DR Pfam: PF01085; HH_signal; 1.
 CC DR PRINTS; PR00632; SONICHHOG.
 CC DR PRODOM; PD003042; HH_signal; 1.
 CC DR SMART; SM00305; HntC; 1.
 CC DR SMART; SM00306; HntN; 1.
 CC DR PROSITE; PS50817; INTERIN_NTER; 1.
 CC KW Developmental protein; Autocatalytic cleavage; Hydrolyase; Protease;
 CC Signal; Lipoprotein; Palmitate; Disease mutation; Holoprosencephaly.
 CC FT SIGNAL 1 23
 CC FT CHAIN 24 462
 CC FT CHAIN 24 197
 CC FT CHAIN 198 462
 CC FT SITE 197 198
 CC FT SITE 243 243
 CC FT SITE 267 267
 CC FT ACT_SITE 270 270
 CC FT BINDING 197 197
 CC FT DOMAIN 407 411
 CC FT LIPID 24 24
 CC FT VARIANT 31 31
 CC FT VARIANT 88 88
 CC FT VARIANT 100 100
 CC FT VARIANT 115 115
 CC FT VARIANT 117 117
 CC FT VARIANT 117 117
 CC FT VARIANT 188 188
 CC FT VARIANT 222 222
 CC FT VARIANT 224 224
 CC FT VARIANT 226 226
 CC FT VARIANT 236 236
 CC FT VARIANT 263 269

/FTId-VAR_009163.
 O -> H (IN HPE: SPORADIC).
 /FTId-VAR_009164.
 N -> K (IN HPE: FAMILIAL).
 /FTId-VAR_009165.
 W -> G (IN HPE).
 /FTId-VAR_003620.
 W -> R (IN HPE).
 /FTId-VAR_003621.
 E -> Q (IN HPE: FAMILIAL).
 /FTId-VAR_009166.
 D -> N (IN HPE: FAMILIAL).
 /FTId-VAR_009167.
 V -> E (IN HPE).
 /FTId-VAR_009168.
 A -> T (IN HPE: FAMILIAL).
 /FTId-VAR_009169.
 S -> R (IN HPE: FAMILIAL).
 /FTId-VAR_009170.
 MISSING (IN HPE: SPORADIC).

Query Match 97.5%; Score 2407; DB 1; Length 462;
 Best Local Similarity 100.0%; Pred. No. 2,4e-170;
 Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVIVSLVSGIACGPGCGKRRHKKLPLVAKOTIPVAKETGASG 60
 DB 1 MLLARCLLVIVSLVSGIACGPGCGKRRHKKLPLVAKOTIPVAKETGASG 60
 QY 61 RYEGKISRNSERKELTPNPNPDIFKENTGADRLMTORCKDKLAISVNMOPGV 120
 DB 61 RYEGKISRNSERKELTPNPNPDIFKENTGADRLMTORCKDKLAISVNMOPGV 120
 QY 121 KLVTEGMDGHHSESLHYEGRAVDITTSDBRSKYGMILARLAVAGEFDWYYESKAH 180
 DB 121 KLVTEGMDGHHSESLHYEGRAVDITTSDBRSKYGMILARLAVAGEFDWYYESKAH 180
 QY 181 IHCSVKAENSVAAKSGCGFCGSAVHLEOGTKLVKLSGDRVLADDDGRLLYSDFLT 240
 DB 181 IHCSVKAENSVAAKSGCGFCGSAVHLEOGTKLVKLSGDRVLADDDGRLLYSDFLT 240
 QY 241 FLDRDCAKVFYIETREPERELLTAHLFLVAPHNSATGEPASGSGPSGALG 300
 DB 241 FLDRDCAKVFYIETREPERELLTAHLFLVAPHNSATGEPASGSGPSGALG 300
 QY 301 PRALFASRVPRGQVYVVAERDGRLLPAVAHSVTLSEBAGAYAPLTAQGTLLINRVL 360
 DB 301 PRALFASRVPRGQVYVVAERDGRLLPAVAHSVTLSEBAGAYAPLTAQGTLLINRVL 360
 QY 361 ASCAVYEESHMAHAPAPRLAHLAALAPARTDNGSGGCGGGRVALTPAGA 420
 DB 361 ASCAVYEESHMAHAPAPRLAHLAALAPARTDNGSGGCGGGRVALTPAGA 420
 QY 421 ADAPGAGATGIMHYSQLTOIGTWLDSALHPILGMAVSS 462
 DB 421 ADAPGAGATGIMHYSQLTOIGTWLDSALHPILGMAVSS 462

RESULT 2
 SHH_MOUSE STANDARD; PRT: 437 AA.
 AC 062226;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sonic hedgehog protein precursor (SHH) (HMG-1).
 GN SHH OR HHGL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=100950;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=94094334; PubMed=7916661;
 RA Echard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,
 RA McMahon J.A., McMahon A.P.;
 RT "Sonic hedgehog, a member of a family of putative signaling
 RT molecules, is implicated in the regulation of CNS polarity.";
 RL Cell 75:1417-1430(1993).
 RN [2]
 RP REVISION TO 122.
 RC STRAIN=C57BL/6J;
 RA McMahon A.P.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
 RX MEDLINE=95236997; PubMed=7720571;
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
 RA Zhao R., Seidman M.F., Fallon J.F., Beachy P.A.;
 RT "Products, genetic linkage and limb patterning activity of a murine
 RT hedgehog gene.";
 RL Development 120:3339-3353(1994).

RA [4]
 RP FUNCTION, AND AUTOPROTEOLYTIC CLEAVAGE.
 RX MEDLINE=95236997; PubMed=7720571;
 RA Roelink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,
 RA Beachy P.A., Jessell T.M.;
 RT "Floor plate and motor neuron induction by different concentrations of
 RT the amino-terminal cleavage product of sonic hedgehog."
 RL Cell 81:445-455(1995).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 34-195.
 RX MEDLINE=96069744; PubMed=7477329;
 RA Hall T.M.T., Porter J.A., Beachy P.A., Leahy D.J.;
 RT "A potential catalytic site revealed by the 1.7-A crystal structure of
 RT the amino-terminal signalling domain of sonic hedgehog.";
 RL Nature 378:212-216(1995).
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
 CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
 CC FLOOR PLATE-AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN A NUMBER OF EMBRYONIC TISSUES
 CC INCLUDING THE NOTOCHORD, VENTRAL NEURAL TUBE, FLOOR PLATE, LUNG
 CC BUD, ZONE OF POLARIZING ACTIVITY AND POSTERIOR DISTAL MESENCHYME
 CC OF LIMBS. IN THE ADULT, EXPRESSED IN LUNG AND NEURAL RETINA.
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE DURING GASTRULATION.
 CC -1- INDUCTION: BY RETINOIC ACID.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X76290; CMA53922.1; -
 DR PDB: 1VHH; 29-JAN-96.
 DR MEROPS: C46.002; -
 DR MGD: MGI:98297; Shh.
 DR InterPro: IPR000320; HH_signal.
 DR InterPro: IPR001767; Hedgehog_hnt.
 DR InterPro: IPR003586; Hedgehog_hntc.
 DR InterPro: IPR003587; Hedgehog_hntn.
 DR InterPro: IPR002203; Intein.
 DR InterPro: IPR001657; Sonichn.
 DR Pfam: PF01079; Hnt; 1.
 DR Pfam: PF01085; HH_signal; 1.
 DR PRINTS: PR00632; SONICHOG.
 DR ProDom: PD003042; HH_signal; 1.
 DR SMART: SM00305; Hntc; 1.

DR SMART: SM00306; Hntn; 1.
 DR PROSITE: PS50817; INTEN_N_TER; 1.
 DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KM Signal; Lipoprotein; Palmitate; 3D-structure.
 FT SIGNAL 1 24
 FT CHAIN 25 437
 FT CHAIN 25 198
 FT CHAIN 199 437
 FT SITE 198 199
 FT SITE 198 199
 FT SITE 244 244
 FT SITE 268 268
 FT ACT_SITE 271 271
 FT BINDING 198 198
 FT DOMAIN 383 387
 FT LIPID 25 25
 FT SEQUENCE 437 AA; 47773 MW; DOB72F08E7860EF CRC64;

Query Match 83.8%; Score 2068; DB 1; Length 437;
 Best Local Similarity 87.4%; Pred. No. 2.3e-145;
 Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

OY 1 MILLARCLLVVSSLLVCSGLACGPGGFGRRRPRKLTPLATYQFIPIVNAEKTGLGASG 60
 DB 2 LLLARCLVLIASSLLVCPGLACGPGGFGRRRPRKLTPLATYQFIPIVNAEKTGLGASG 61
 OY 61 RYEGKISNRSERFKELTPNYNDPIFKDEENTGADRLMTORCKDKLNLALISVMNQMGV 120
 DB 62 REGKITNSERFKELTPNYNDPIFKDEENTGADRLMTORCKDKLNLALISVMNQMGV 121
 OY 121 KIRTEGDEDEGHSESLHREGRAVDITTSDRDSKYGMLARLAVEAGFDMVYESKAH 180
 DB 122 KIRTEGDEDEGHSESLHREGRAVDITTSDRDSKYGMLARLAVEAGFDMVYESKAH 181
 OY 181 IHCYSKANSVAAGSGGPGSATVHLGOGKTKVLDSPDGRVLAADOGRLYSDPLT 240
 DB 182 IHCYSKANSVAAGSGGPGSATVHLGOGKTKVLDSPDGRVLAADOGRLYSDPLT 241
 OY 241 FLDRDGAKKVYVETETEPRERLTLTAHLFLVPHNDSATGEPEASSGSGPFGSALG 300
 DB 242 FLDRDGAKKVYVETETEPRERLTLTAHLFLVPHNDSATGEPEASSGSGPFGSALG 300
 OY 301 PRALEASVRRGQRYVVAERDGRRLLPAAVHSTVLSSEAGAVAPLTAGCTILINRVL 360
 DB 287 PSALFASVRRGQRYVVAERDGRRLLPAAVHSTVLSSEAGAVAPLTAGCTILINRVL 346
 OY 361 ASCYAVIEHSHARAFAPFLAHALLAALAPARTDGGSGGGRGGGRVALTAPCA 420
 DB 347 ASCYAVIEHSHARAFAPFLAHALLAALAPARTDGGSGGGRGGGRVALTAPCA 395
 OY 421 ADAPGAGTAGIHWSQLLYOIGTWLDSSEALHPLGMAVKSS 462
 DB 396 TEARGAEPTAGIHWSQLLYOIGTWLDSSEALHPLGMAVKSS 437

RESULT 3
 SHH_RAT STANDARD; PRT; 437 AA.
 AC 063673;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sonic hedgehog protein precursor (SHH).
 GN SHH OR VHH-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Embryonic floor plate;

RX MEDLINE-94170375; PubMed-8124714;
 RA Roelink H., Augsburg J., Heemskerk J., Korch V., Norlin S.,
 RA Ruiz J., Altaba A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,
 RA Dodd J.;
 RT "Floor plate and motor neuron induction by vhh-1, a vertebrate homolog
 of hedgehog expressed by the notochord.";
 RL Cell 76:761-775(1994).
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
 CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
 CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE NODE, NOTOCHORD, FLOOR PLATE,
 CC AND POSTERIOR LIMB BUD MESENCHYME.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L27340; AAA20999.1; -
 CC HSSP: 062226; 1VHH.
 DR MEROPS: C46.002; -
 DR InterPro: IPR000320; HH_signal.
 DR InterPro: IPR001767; Hedgehog_hnt.
 DR InterPro: IPR003586; Hedgehog_hntC.
 DR InterPro: IPR003587; Hedgehog_hntN.
 DR InterPro: IPR002203; Intein.
 DR InterPro: IPR001657; SontCHH.
 DR Pfam: PF01079; Hntc; 1.
 DR Pfam: PF01085; HH_signal; 1.
 DR PRINTS: PR00632; SONICHOG.
 DR PRODOM: PD003042; HH_signal; 1.
 DR SMART: SM00305; Hntc; 1.
 DR SMART: SM00306; Hntn; 1.
 DR PROSITE: PS50817; INTEN_N_TER; 1.
 DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KM Signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 24
 FT CHAIN 25 437
 FT CHAIN 25 198
 FT CHAIN 199 437
 FT SITE 198 199
 FT SITE 198 199
 FT SITE 244 244
 FT SITE 268 268
 FT ACT_SITE 271 271

FT BINDING 198 198 SIMILARITY).
 FT DOMAIN 383 387 CHOLESTEROL (BY SIMILARITY).
 FT LIPID 25 25 POLY-GLY.
 SQ SEQUENCE 437 AA; 47630 MW; 0DBEC19F0D1662A0 CRC64;
 Query Match 83.4%; Score 2060; DB 1; Length 437;
 Best Local Similarity 87.5%; Pred. No. 9.2e-145;
 Matches 405; Conservative 8; Mismatches 22; Indels 28; Gaps 4;

QY 1 MLLALCLLVVSSILVCSGLACGGRGRRHPRKLLPLVAKQFIPVNAKTTGASG 60
 DB 2 LLLALCFVLAASLLVCPGLACGPRGRGKRQPKLLPLVAKQFIPVNAKTTGASG 61
 QY 61 RYEGKISRNSERKELTPNPNDIIFKEENTGADRLMTORCKDKLNALISYANQMPGV 120
 DB 62 RYEGKITRNSERKELTPNPNDIIFKEENTGADRLMTORCKDKLNALISYANQMPGV 121
 QY 121 KLAVTEGMDGHHSESLHYEGRAVDITTSDBDRSKYGMARLAVEAGFDWVYVESKAR 180
 DB 122 KLAVTEGMDGHHSESLHYEGRAVDITTSDBDRSKYGMARLAVEAGFDWVYVESKAR 181
 QY 181 IHSVVAENVAKSGGCFGSGATVHLEGGTKLVNDLSFGDRLVLAADDGRLYSDFLT 240
 DB 182 IHSVVAENVAKSGGCFGSGATVHLEGGTKLVNDLSFGDRLVLAADDGRLYSDFLT 241
 QY 241 FLDRDQAKVYFYIETREPERELLTAHLLEVAAPHNDATGPEASGSGPPSGALG 300
 DB 242 FLDRDQAKVYFYIETREPERELLTAHLLEVAAPHNDATGPEASGSGPPSGALG 300
 QY 301 PRAFLASRVPRGQRYVVAERDRLPLPAVHVSVTLSSEBAGAYALTAOGTLLINRVL 360
 DB 287 PSFLFASRVPRGQRYVVAERDRLPLPAVHVSVTLSSEBAGAYALTAOGTLLINRVL 346
 QY 361 ASCVAYVEESHMAHRAAPRLAHALLAALAPARTDGGSGGGDRGGGGRVALTAP-G 419
 DB 347 ASCVAYVEESHMAHRAAPRLAHALLAALAPARTDGGSGGGDRGGGGRVALTAP-G 394
 QY 420 AADAPGAGATAGIHWSQLLYOIGTWLIDSEALHPLGMAYKSS 462
 DB 395 VAERAGSGPPAGIHWSQLLYOIGTWLIDSEALHPLGMAYKSS 437

RESULT 4
 SHH_CHICK STANDARD; PRT; 425 AA.
 AC 091035;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sonic hedgehog protein precursor (SHH).
 GN SHH.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 ON NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=limb bud;
 RX MEDLINE=94094333; PubMed=8269518;
 RA Riddle R.D., Johnson R.L., Lauffer E., Tabin C.,
 RA "Sonic hedgehog mediates the polarizing activity of the ZPA."
 RL Cell 75:1401-1416(1993).
 RN [2]
 RP FUNCTION AND AUTOPROTEOLYTIC CLEAVAGE.
 RX MEDLINE=95254654; PubMed=7736596;
 RA Roelink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,
 RA Beachy P.A., Jessell T.M.;
 RT "Floor plate and motor neuron induction by different concentrations of
 RT the amino-terminal cleavage product of sonic hedgehog
 RT autoproteolysis".
 RL Cell 81:445-455(1995).

CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
 CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESENTS THE
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
 CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE POSTERIOR LIMB BUD
 CC MESENCHYME, THE HENSEN'S NODE, THE NOTOCHORD, AND THE FLOOR PLATE
 CC OF THE NEURAL TUBE.
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT STAGE 17 DURING THE
 CC INITIATION OF LIMB BUD FORMATION. FROM THAT POINT ONWARDS, THE
 CC EXPRESSION PATTERN EXACTLY MATCHES THE LOCATION OF THE ZONE OF
 CC POLARIZING ACTIVITY (ZPA).
 CC -1- INDUCTION: BY RETINOID ACID.
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L28099; AAA72428.1; -
 CC HSSP: O62226; 1VHH.
 CC DR MEROPS: C46.001; -
 CC DR InterPro: IPR000320; HH_signal.
 CC DR InterPro: IPR001767; Hedgehog_hint.
 CC DR InterPro: IPR003586; Hedgehog_hintc.
 CC DR InterPro: IPR003587; Hedgehog_hintn.
 CC DR InterPro: IPR002203; Intein.
 CC DR InterPro: IPR001657; SonICHH.
 CC Pfam: PF01079; Hint; 1.
 CC Pfam: PF01085; HH_signal; 1.
 CC PRINTS: PR00632; SONICHOG.
 CC PRODOM: PD003042; HH_signal; 1.
 CC SMART: SM00305; Hintc; 1.
 CC SMART: SM00306; Hintn; 1.
 CC PROSITE: PS50817; INTEIN_N_TER; 1.
 CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 CC Signal; Lipoprotein; Palmitate.
 CC FT SIGNAL 1 26
 CC FT CHAIN 27 425
 CC FT CHAIN 27 200
 CC FT CHAIN 201 425
 CC FT SITE 200 201
 CC FT SITE 246 246
 CC FT SITE 270 270
 CC FT SITE 273 273
 CC ACT_SITE 273 273
 CC BINDING 200 200
 CC
 CC SONIC HEDGEHOG PROTEIN.
 CC SONIC HEDGEHOG PROTEIN N-PRODUCT.
 CC SONIC HEDGEHOG PROTEIN C-PRODUCT.
 CC CLEAVAGE (AUTO-) (PROBABLE).
 CC INVOLVED IN CHOLESTEROL TRANSFER (BY
 CC SIMILARITY).
 CC INVOLVED IN AUTO-CLEAVAGE (BY
 CC SIMILARITY).
 CC ESSENTIAL FOR AUTO-CLEAVAGE (BY
 CC SIMILARITY).
 CC CHOLESTEROL (BY SIMILARITY).

FT DOMAIN 390 393 POLY-THR.
 FT LIPID 27 27 PALMITATE (BY SIMILARITY).
 SQ SEQUENCE 425 AA; 46474 MM; DA9627443DA0173 CRC64;

Query Match 73.2%; Score 1807; DB 1; Length 425;
 Best local similarity 77.8%; Pred. No. 4,1e-126;
 Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;

OY 1 MLLARCLLVVSLVCSGLACGPGRGFRKRRPKLTPLAYKQFIPNVAEKTIGASG 60
 DB 4 MLLRLRLLVGFICALVSSGLTCGPGRGIGRRRPRKLTPLAYKQFIPNVAEKTIGASG 63
 OY 61 RYEGKISNSRFRKLTPLNPNYNDIIFKDEENTGADRLTORCKKLNLAISVNMQPGV 120
 DB 64 RYEGKITNSRFRKLTPLNPNYNDIIFKDEENTGADRLTORCKKLNLAISVNMQPGV 123
 OY 121 KLRTVGEDEGHHSESLHTEGRAVDITTSDRSKYGMRLAVEGFPWVYESKAH 180
 DB 124 KLRTVGEDEGHHSESLHTEGRAVDITTSDRSKYGMRLAVEGFPWVYESKAH 183
 OY 181 IHCYKAENSVAAKSGCGFPGSATVHLEOGGTKLVDLSPGDRVLAADGRLKLYSDFLT 240
 DB 184 IHCYKAENSVAAKSGCGFPGSATVHLEOGGTKLVDLSPGDRVLAADGRLKLYSDFLT 243
 OY 241 FLDRDDGAKKVFYVETREPRRLTLTAHLLEVPAR-HNDS-ANGEPPASSGSGPPSGCA 298
 DB 244 FLDRDDGAKKVFYVETREPRRLTLTAHLLEVPAR-HNDS-ANGEPPASSGSGPPSGCA 292
 OY 299 LGPRALFASRVPRGORYVVAERDGRRLTPAAVHSVTLSEBAAGAVAPLTAQGTILINR 358
 DB 293 ---QALFASNVKPGORYVVLGE--GGQOLPLSVHSVSLREASGAVAPLTAQGTILINR 347
 OY 359 VLASCVAVIEHSMHRAFAFRLNALALAPARTDRGDSGGGGRVALTPAL 418
 DB 348 VLASCVAVIEHSMHRAFAFRLNALALAPARTDRGDSGGGGRVALTPAL 418
 OY 419 GAADAPGAGATAGTHTWSQLYQIGTWTLLDSEALHPICMAVAKSS 462
 DB 382 DGALPTAATTTTGHWSRLYRIGSVLDDGDLHPICMAVAPAS 425

RESULT 5
 ID SHH_CYNPY STANDARD; PRT; 432 AA.
 AC 090385;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sonic hedgehog protein precursor (SHH).
 GN SHH.
 OS Cynops pyrrhogaster (Japanese common newt).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Caudata; Salamandridae; Salamandridae; Cynops.
 OX NCBI_TaxID=8330;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Embryo;
 RX MEDLINE=96136334; PubMed=8573168;
 RA Takahashi T., Takahashi T.C., Inoue K., Ogawa M., Takashina K.;
 RT "Activation of two Cynops genes, fork head and sonic hedgehog, in
 RL animal cap explants.";
 RL Biochem. Biophys. Res. Commun. 218:395-401(1996).
 CC 1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED
 CC IN LIMB FORMATION, PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND
 CC VENTRAL SOMITE DIFFERENTIATION. INDUCES ECTOPIC CEMENT GLAND
 CC FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH
 CC TRANSDUCES SIGNALS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE
 CC TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC
 CC REPRESENTS THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE

CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- INDUCTION: ACTIVATED BY ACTININ, BASIC FIBROBLAST GROWTH FACTOR
 CC (BFGF) AND FORK HEAD.
 CC -1- PRT: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).

DR EMBL; D63339; BAA09657.1; .
 DR HSSP; Q62226; 1VHH.
 DR MEROPS; C46.002; .
 DR InterPro: IPR000320; HH_signal.
 DR InterPro: IPR001767; Hedgehog_hnt.
 DR InterPro: IPR003586; Hedgehog_hntc.
 DR InterPro: IPR003587; Hedgehog_hntcn.
 DR InterPro: IPR002203; Intein.
 DR InterPro: IPR001657; SonicHH.
 DR Pfam: PF01079; Hnt; 1.
 DR Pfam: PF01085; HH_signal; 1.
 DR PRINTS; PR00632; SONICHHOG.
 DR ProDom; PD003042; HH_signal; 1.
 DR SMART; SM00305; Hntc; 1.
 DR SMART; SM00306; Hntcn; 1.
 DR PROSITE; PS50817; INTEIN_N_TER; 1.
 DR Developmental protein: Autocatalytic cleavage; Hydrolyase; Protease;
 KW Signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 26
 FT CHAIN 27 432
 FT CHAIN 27 200
 FT CHAIN 201 432
 FT SITE 200 201
 FT SITE 268 268
 FT ACT_SITE 271 271
 FT BINDING 200 200
 FT LIPID 27 27
 FT LIPID 27 27
 SQ SEQUENCE 432 AA; 47847 MM; B455C7E746C8E5A8 CRC64;

Query Match 62.4%; Score 1539.5; DB 1; Length 432;
 Best local similarity 67.7%; Pred. No. 2.3e-106;
 Matches 315; Conservative 36; Mismatches 75; Indels 39; Gaps 8;

OY 1 MLLARCLLVVSLVCSGLACGPGRGFRKRRPKLTPLAYKQFIPNVAEKTIGASG 60
 DB 4 MLLRLRLLVGFICALVSSGLTCGPGRGIGRRRPRKLTPLAYKQFIPNVAEKTIGASG 63
 OY 61 RYEGKISNSRFRKLTPLNPNYNDIIFKDEENTGADRLTORCKKLNLAISVNMQPGV 120
 DB 64 RYEGKITNSRFRKLTPLNPNYNDIIFKDEENTGADRLTORCKKLNLAISVNMQPGV 123
 OY 121 KLRTVGEDEGHHSESLHTEGRAVDITTSDRSKYGMRLAVEGFPWVYESKAH 180
 DB 124 KLRTVGEDEGHHSESLHTEGRAVDITTSDRSKYGMRLAVEGFPWVYESKAH 183
 OY 181 IHCYKAENSVAAKSGCGFPGSATVHLEOGGTKLVDLSPGDRVLAADGRLKLYSDFLT 240
 DB 184 IHCYKAENSVAAKSGCGFPGSATVHLEOGGTKLVDLSPGDRVLAADGRLKLYSDFLT 243

```

OY 241 FLDRODGAKKVVYETREPRERLTLTAALLFVA---PHNDATGEPEASSGSGPPSGG 297
DB 244 FMDKEETKRVFYETI--SREVRLLTAHLLEFVGAHGNND-----SGGDFRSVFG 293
OY 298 ALCPRALFASRRVPCGRVYVVAERODDRLLPAAVHSVYLSEBAAGAYAPLTAOGTILIN 357
DB 294 SAGFRSMFASVAGHGRVLTV-DREG-RGLREATVERVYL-EEATGAYAPVTAHGTIVID 350
OY 358 RVLASCYATIEEHSMAHRAFAPERLAHLAALAPARTRGDGGSGGGGGRVALTA 417
DB 351 RVLASCYATIEEHSMAHRAFAPERLAHLAALAPARTRGDGGSGGGGGRVALTA 417
OY 418 PGAAADPGAGATAGIHMYSQLYQIGTWLIDSEALHPILGMAVKS 462
DB 393 -----PPAPSGEGVHMYSIELYRIGTWTWLOEDTIIHPLGMAKSS 432

RESULT 6
SHH_BRARE
ID SHH_BRARE STANDARD; PRT; 418 AA.
AC 092008; 013170; 013171;
DT 15-JUL-1999 (rel. 38, Created)
DT 15-JUL-1999 (rel. 38, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Sonic hedgehog protein precursor (SHH) (VHH-1).
GN SHH OR VHH1.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
[1]
RN RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE-94170375; PubMed-8124714;
RA Koelink H., Augsburg A., Heemskerk J., Korzh V., Norlin S.,
RA Ruidz I., Altaba A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,
RA Dodd J.;
RT "Floor plate and motor neuron induction by vhh-1, a vertebrate homolog
of hedgehog expressed by the notochord.";
RL Cell 76:761-775(1994).
[2]
RN RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
RX MEDLINE-96014264; PubMed-7583153;
RA Ekker S.C., Ungar A.R., Greenstein P., von Kessler D.P., Porter J.A.,
RA Moon R.T., Beachy P.A.;
RT "Patterning activities of vertebrate hedgehog proteins in the
developing eye and brain.";
RL Curr. Biol. 5:944-955(1995).
[3]
RN RP SEQUENCE FROM N.A.
RX MEDLINE-96083328; PubMed-7579523;
RA Fietz M.J., Concorde J.-P., Barbosa R., Johnson R., Krauss S.,
RA McMahon A.P., Tabin C., Ingham P.W.;
RT "The hedgehog gene family in Drosophila and vertebrate development.";
RL Development Suppl. 43-51(1994).
[4]
RN RP SEQUENCE FROM N.A.
RX MEDLINE-99225458; PubMed-10207136;
RA Muller F., Chang B., Albert S., Fischer N., Tora L., Strahle U.;
RT "Intronic enhancers control expression of zebrafish sonic hedgehog in
floor plate and notochord.";
RL Development 126:2103-2116(1999).
[5]
RN RP SEQUENCE OF 30-92 AND 113-170 FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE-97075114; PubMed-8917540;
RA Zarolaya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
closely related to the zebrafish.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
-1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF

```

```

CC CC PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
CC CC NOTOCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING
CC CC OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS
CC CC FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC)
CC CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO
CC CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,
CC CC PTC REPRESSED THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
CC CC SIMILARITY).
-1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
-1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL
CC CC TUBE AND BRAIN. ALSO FOUND IN THE NOTOCHORD AND IN DEVELOPING FIN
CC CC BUD. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN DOMAINS THAT
CC CC INCLUDE A DISCRETE REGION IN THE FLOOR OF THE DIENCEPHALON.
-1- DEVELOPMENTAL STAGE: FIRST DETECTABLE IN THE INNER CELL LAYER OF
CC CC THE EMBRYONIC SHIELD DURING GASTRULATION. BY 9.5 HRS OF
CC CC DEVELOPMENT, EXPRESSED IN A CONTINUOUS BAND THAT EXTENDS FROM THE
CC CC TAIL TO THE HEAD, THE ANTERIOR BOUNDARY OF EXPRESSION BEING
CC CC POSITIONED IN THE CENTER OF THE ANIMAL POLE ANTERIOR TO THE
CC CC PRESUMPTIVE MIDBRAIN.
-1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
-1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC CC or send an email to license@slb-slb.ch).
-----
CC CC EMBL; L27585; AAA20998.1; -
CC CC EMBL; U30711; AAC59742.1; -
CC CC EMBL; Z35669; CA84738.1; -
CC CC EMBL; AF124382; AAD47913.1; -
CC CC EMBL; U51351; AAB38575.1; -
CC CC EMBL; U51370; AAB38593.1; -
CC CC HSP; Q62226; 1VHH.
CC CC MEROPS; C46.002; -.
CC CC ZFIN; ZDB-GENE-980526-166; shh.
CC CC InterPro: IPR000320; HH_signal.
CC CC InterPro: IPR001767; Hedgehog_hhnt.
CC CC InterPro: IPR003586; Hedgehog_hhntc.
CC CC InterPro: IPR003587; Hedgehog_hhntn.
CC CC InterPro: IPR002203; Intein.
CC CC InterPro: IPR001657; SonicHH.
CC CC Pfam; PF01079; Hhnt; 1.
CC CC Pfam; PF01085; HH_signal; 1.
CC CC PRINTS; PR00632; SONICHHOG.
CC CC PRODOM; PD003042; HH_signal; 1.
CC CC SMART; SM00305; Hhntc; 1.
CC CC SMART; SM00306; Hhntn; 1.
CC CC PROSITE; PS50817; INTEIN_N_TER; 1.
CC CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
CC CC signal; Lipoprotein; Palmitate.
CC CC SIGNAL; 1; 23
CC CC CHAIN; 24; 418
CC CC CHAIN; 24; 197
CC CC CHAIN; 198; 418
CC CC SITE; 197; 198
CC CC SITE; 243; 243
CC CC SITE; 267; 267
CC CC INVOLVED IN AUTO-CLEAVAGE (BY
CC CC SIMILARITY).

```



```

FT CHAIN 25 444 SONIC HEDGEHOG PROTEIN,
FT CHAIN 25 198 SONIC HEDGEHOG PROTEIN N-PRODUCT,
FT CHAIN 199 444 SONIC HEDGEHOG PROTEIN C-PRODUCT,
FT SITE 198 199 CLEAVAGE (AUTO-) (BY SIMILARITY),
FT SITE 266 266 INVOLVED IN AUTO-CLEAVAGE (BY
FT ACT_SITE 269 269 SIMILARITY),
FT DOMAIN 386 409 ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT REPEAT 386 393 3 X 8 AA TANDEM REPEATS OF Q-V-D-L-Q-S-H-
FT REPEAT 394 401 1.
FT REPEAT 403 409 2.
FT BINDING 198 198 CHOLESTEROL (BY SIMILARITY),
FT LIPID 25 25 PALMITATE (BY SIMILARITY),
FT CONFLICT 5 9 TOSIL -> NSMLCW (IN REF. 3),
FT CONFLICT 302 319 DPKMTKAVKEKVD -> ESDHDHLEGKMKRLLR
FT CONFLICT 432 432 (IN REF. 3),
FT CONFLICT 432 432 N -> S (IN REF. 3),
FT SEQUENCE 444 AA: 49453 MW: 7384E932FA2EFF2 CRC64:

Query Match 60.7%: Score 1498; DB 1; Length 444;
Best Local Similarity 64.1%: Pred. No. 2,8e-103;
Matches 305; Conservative 38; Mismatches 87; Indels 46; Gaps 7;

OY 1 MLLARCLLY-LVSSLVCSGLACGPGRGGRHRRPKLPLAYKQFIIPNAEKTIGAS 59
DB 1 MLYATQSLLESLFICLTLPPLGACGPGRGGRHRRPKLPLAYKQFIIPNAEKTIGAS 60
OY 60 GVEEKISNSRFKELTPNYPDIIFDEENTGADRLTQCKCKLNLAISSNMQMG 119
DB 61 GVEEKITNSKCFELTPNYPDIIFDEESTGADRLTQCKCKLNLAISSNMQMG 120
OY 120 VALRTGEGDEGHSSESLAYEGRAVDITTSDRDSKSGMLARLAVEGPMVYESKA 179
DB 121 VALRTGEGDEGHSSESLAYEGRAVDITTSDRDSKSGMLARLAVEGPMVYESKA 180
OY 180 HHSVKAENSVAAKSGCGPGSATVHLEQGGTKLVKDISPDRLVLAADGGRLYSDFL 239
DB 181 HHSVKAENSVAAKSGCGPGSATVHLEQGGTKLVKDISPDRLVLAADGGRLYSDFL 240
OY 240 TFLDDDDAKKFFVYIEPEPRERLLTAAHLFFAAPHNDSATGEPASSSGPPGGL 299
DB 241 MTDIDRDVKKFFVYIEPEPRERLLTAAHLFFAAPHNDSATGEPASSSGPPGGL 282
OY 300 GPRALFASRPVPGQRYVVAERDCDRRLPAVHSTVLSSEAAYAPLTAOCTILINIV 359
DB 283 SKKSVFASRIQPGDLITADPK--TMTLKAVYKEVDL-EEDTGAYAPLTAOCTIVIDQY 339
OY 360 LASCYAVIEHSMARAFAPRLAHLAALAPARTDRGDSGGDRGGGG----- 411
DB 340 LASCYAVIEHSMARAFAPRLAHLAALAPARTDRGDSGGDRGGGG----- 411
OY 412 -----RVALTARGADAPAGATAGIHWSQLLYQGTGLDSEALHPRGMAVKS 462
DB 389 LOSHNOVDLQSHNOVDLQSHNOVDLQSHNOVDLQSHNOVDLQSHNOVDLQSHNOVDL 444

RESULT 8
TMHH_BRAE
ID TMHH_BRAE STANDARD: PRT: 416 AA.
AC Q90419.
DT 15-JUL-1999 (rel. 38, Created)
DT 15-JUL-1999 (rel. 38, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE TIGGY-WINKLE hedgehog protein precursor (TMHH).
GN TMHH.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]

```

```

RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
RC TISSUE=Embryo;
RX MEDLINE=96014264; PubMed=7583153;
RA Eker S.C., Ungar A.R., Greenstein P., von Kessler D.P.,
RA Porter J.A., Moon R.T., Beachy P.A.;
RT "Patterning activities of vertebrate hedgehog proteins in the
RT developing eye and brain.";
RL Curr. Biol. 5:944-955(1995).
CC -!- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
CC PATTERNING EVENTS DURING DEVELOPMENT. INVOLVED IN DORSO-VENTRAL
CC PATTERNING OF THE BRAIN AND IN EARLY PATTERNING OF THE DEVELOPING
CC EYES.
CC -!- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL
CC TUBE AND BRAIN. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN
CC DOMAINS THAT INCLUDE A DISCRETE REGION IN THE FLOOR OF THE
CC DIENCEPHALON. NOT DETECTED IN THE NOTOCHORD OR DEVELOPING FIN BUD.
CC -!- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC
CC DR HSMB: U30710; AAC59741.1; -.
CC DR HSMB: Q62226; VAMH.
CC DR MEROPS: C46.001; -.
CC DR ZFIN: ZDB-GENE-960526-41; twhh.
CC DR InterPro: IPR000320; HH_signal.
CC DR InterPro: IPR001767; Hedgehog_hntc.
CC DR InterPro: IPR003586; Hedgehog_hntc.
CC DR InterPro: IPR003587; Hedgehog_hntn.
CC DR InterPro: IPR001657; Son1cHh.
CC DR Pfam: PF01079; Hnt; 1.
CC DR Pfam: PF01085; HH_signal; 1.
CC DR PRINTS: PR00632; SONICHHOG.
CC DR ProDom: PD003042; HH_signal; 1.
CC DR SMART: SM00305; Hntc; 1.
CC DR SMART: SM00306; Hntn; 1.
CC DR SMART: SM00305; Hntc; 1.
CC DR SMART: SM00306; Hntn; 1.
CC KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
CC signal; Lipoprotein; Palmitate.
CC
CC FT SIGNAL 1 26
CC FT CHAIN 27 416
CC FT CHAIN 27 200
CC FT CHAIN 201 416
CC FT SITE 200 201
CC FT SITE 270 270
CC FT ACT_SITE 273 273
CC FT BINDING 200 200
CC FT LIPID 27 27
CC FT SEQUENCE 416 AA: 46576 MW: 61EC2218309CE59 CRC64:

Query Match 58.0%: Score 1432; DB 1; Length 416;
Best Local Similarity 62.6%: Pred. No. 1,9e-98;
Matches 288; Conservative 39; Mismatches 77; Indels 56; Gaps 5;

```

QY	2	LLARCLLLVSSLLVSSGLACGSPRGKGRHKRPKLLPVLAKQFIPIVNAEKTGASGR	61
QY	122	LRVTEGNDDEGHHSEELHTEGAVALITTSDDRKSYGGGLAFLAVLAGEFDWYIESKAHI	181
QY	125	LRVTEGNDDEGHHLESLHTEGAVALITTSDDRKSYGGGLAFLAVLAGEFDWYIESKAHI	184
QY	182	HCSYKAENSVAAKSGGCPSPATVTHLEGGGTIKYVKOLSPGDVTLAADOGRLLTSDPTE	241
QY	185	HCSYKAENSVAAKSGGCPSPATVTHLEGGGTIKYVKOLSPGDVTLAADOGRLLTSDPTE	244
QY	242	LDRDGAKKVFYVLETPERPERLLTAAHLFLVAPRANSDATPEPAASSGSGPPSGALCP	301
QY	245	IDHDPTTRQFIYETIESEPTTKTLTNAHLVTVG--NSAA-----ASGI	286
QY	302	RALFASVNRPGORVYVAERDGRKLLPAVHVSVTL---SEEAQAVALPLTAOGTILIN	357
QY	287	TATFASVNRPGORVYVAERDGRKLLPAVHVSVTL---SEEAQAVALPLTAOGTILIN	339
QY	358	RVLSCAVALTEESHNRAPARPLRALLAALAPRDRDGGDSGGGRGGGGRVALTA	417
QY	340	QVLASCAVALTEESHNRAPARPLRALLAALAPRDRDGGDSGGGRGGGGRVALTA	383
QY	418	PGAADAPGATAGTGINHYSQLTQITGWTLLDSEALPIGLM	457
QY	384	-----EDGIIHYSNMLFHTISWLLDSDSPHPIGI	412
RESULT	9		
INH_CHICK	ID	INH_CHICK	STANDARD: PRT: 408 AA.
AC	098938:		
DT	15-JUL-1999	(Rel. 38, Created)	
DT	15-JUL-1999	(Rel. 38, Last sequence update)	
DT	16-OCT-2001	(Rel. 40, Last annotation update)	
DE	Indian hedgehog protein precursor (INH).		
GN	gallus gallus (Chicken).		
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
OC	Gallus.		
OX	NCBI_TaxID=9031;		
RN	RP	SEQUENCE FROM N.A.	
RA	MEDLINE=6632546;		
RA	Vortkamp A., Lee K., Lanske B., Segre G.V., Kronenberg H.M.,		
RA	Tabin C.J.;		
RT	"Regulation of rate of cartilage differentiation by Indian hedgehog		
RT	and PTH-related protein."		
RT	Science 273:613-622(1996).		
CC	-1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF		
CC	PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)		
CC	RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO		
CC	ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN		
CC	ENDOCHONDRAL OSSIFICATION: MAY REGULATE THE BALANCE BETWEEN GROWTH		
CC	AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION		
CC	OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP).		
CC	-1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE		
CC	CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE		
CC	CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM		
CC	TO MEDIAN SIGNALING TO OTHER CELLS (BY SIMILARITY).		
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN DEVELOPING MIDGUT, LUNG AND		
CC	CARTILAGE OF DEVELOPING LONG BONES IN THE LIMB.		
CC	-1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY		
CC	AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN		
CC	THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF		
CC	A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-		
CC	TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS		

CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@lsb.slb.ch).

DR EMBL; U58511; AAC60010.1; -.
DR HSSP; Q62226; IVHH.
DR MEROPS; C46_003; -.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001767; Hedgehog_hntc.
DR InterPro; IPR003586; Hedgehog_hntc.
DR InterPro; IPR003587; Hedgehog_hntn.
DR InterPro; IPR001657; Sonichh.
DR Pfam; PF01079; Hnt; 1.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHG.
DR ProDom; PD003042; HH_signal; 1.
DR SMART; SMO0305; HntC; 1.
DR SMART; SMO0306; Hntn; 1.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KM Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 23
FT CHAIN 24 408
FT CHAIN 24 198
FT CHAIN 199 408
FT SITE 198 199
FT SITE 244 244
FT SITE 268 268
FT ACT_SITE 271 271
FT BINDING 198 198
FT LIPID 24 24
SQ SEQUENCE 408 AA; 44829 MM; BA397AE2A9357A24 CRC64;

Query Match 51.9%; Score 1282.5; DB 1; Length 408;
Best Local Similarity 59.7%; Pred. No. 2e-87;
Matches 271; Conservative 34; Mismatches 92; Indels 57; Gaps 8;

QY 5 ARCLLLVYSSLVCSGL-AAGPGRGFG-KRRHPKLTPLAYKQFIPNVAEKTLAGSGRY 62
Db ||||| :||: ||||| :|||| |:||:||||||| ||||| |||||
4 ARLLLLLGCALLLPAAVRCGPGRVGVSSRRRPKKLLPLAVKQSPVPEKTLGASGRY 63

DY . 63 EGKSIRNSERKELTPTNPNDIIFFDEENTGADRLMTORCKDKLNALAIYSNNOPGYKL 122
Db EGTARNSEKERELTPNTNPDIIFDEENTGADRLMTORCKDLNLSLAISYNNOPGYKL 123

QY 123 RVTGMDDGDGHSEESLIEGRAVDITTSDDRKYGGMLARLAYAGGDWYYESKAHIH 182
Db RTVEGMDDGDGHSEESLIEGRAVDITTSDDRKNYGGMLARAAYAGDWMYYESKAHIH 183

QY 183 CSYAENSVAAKSGCGPFSATVHLEOGTRLYVDLSFGDRVLAADDGRLIYSDFLFL 242
Db CSVKSEHSAAKKGTCGCFGALATLENGAFRPLMALRGQGVYLMAMDGAGRPTYSDFLFL 243

QY 243 DRDGAARKVFIVITREPREKRLTLTAHLLVAPHNDSATGEPEASSGSGPPSGGALCP 302
Db DKEPRALTAFVETROPRLRALTPHTLFEVA---DNA-----SAFAOFR 287

QY 303 ALFASPVAGGRVVVAREDDRRLLPLPAVSYLTSEBAAGAVPLTAOGTILIMRYVAS 362
Db PTFASHVQPGHFVLVAV---GSGGLQPLEVVGVR-GRTDVGAIAPLATHGLVYVDVYVAS 343


```

Db 350 AVADHHHQAFLAFMPLRLRHSLSL-----AMGSPWG 378
Oy 426 AGATAGCCHWYSQLLYQYGTGTHLLDSEALHPLGNAWSSXSRGAG 468
Db 379 -----EGVHWYRQQLLYRLGRLLRLEEGSGFHLGMS-----GAG 410

RESULT 11
ID IHH_MOUSE STANDARD: PRT; 411 AA.
AC P97812; 061724;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Indian hedgehog protein precursor (IHH) (Hh-2).
GN IHH.
OS Mus musculus (Mouse).
OC Eumetazoa; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
RC STRAIN=CD-1; TISSUE=Kidney;
RX MEDLINE=97236802; PubMed=9079674;
RA Vaeshtini R.P., Brookhiser W.T., Park J., Yang T., Briggs J.,
RA Dressler G., Holzman L.B.;
RT "Post-translational processing and renal expression of mouse Indian
RT hedgehog.";
RL J. Biol. Chem. 272:8466-8473(1997).
RN 12
RP SEQUENCE OF 76-411 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=embryo;
RX MEDLINE=94094334; PubMed=7916661;
RA Echejard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,
RA McMahon J.A., McMahon A.P.;
RT "Sonic hedgehog, a member of a family of putative signaling
RT molecules, is implicated in the regulation of CNS polarity.";
RL Cell 75:1417-1430(1993).
RN 13
RP REVISONS.
RC STRAIN=C57BL/6J;
RA St Jacques B.;
RL Submitted (APR-1997) to the EMBL/Genbank/DBJ databases.
RN 14
RP SEQUENCE OF 124-172 FROM N.A.
RX MEDLINE=95236997; PubMed=7720571;
RA Chang D.T., Lopez A., von Kesseler D.P., Chiang C., Simandl B.K.,
RA Zhao R., Seidlin M.F., Fallon J.F., Beachy P.A.;
RT "Products, genetic linkage and limb patterning activity of a murine
RT hedgehog gene.";
RL Development 120:3339-3353(1994).
RN 15
RP FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
RP PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)
RP RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO
RP ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN
RP ENDOCHONDRAL OSSIFICATION: MAY REGULATE THE BALANCE BETWEEN GROWTH
RP AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION
RP OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP).
RN 16
RP SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
RP CELL SURFACE. WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
RP CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
RP TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
RN 17
RP TISSUE SPECIFICITY: IN THE EMBRYO. DETECTED IN THE DEVELOPING GUT,
RP THE GROWTH ZONE OF CARTILAGE OF DEVELOPING LONG BONES, EPITHELIUM
RP AND UROGENITAL SINUS. IN THE ADULT KIDNEY, FOUND IN PROXIMAL
RP CONVOLUTED AND PROXIMAL STRAIGHT TUBULE.
RN 18
RP DEVELOPMENTAL STAGE: DETECTED AT 10 DAYS POST COITUM (DPC) IN
RP THE DEVELOPING GUT, AT 14.5 DAYS DPC IN THE CARTILAGE PRIMORDIUM AND
RP IN THE DEVELOPING UROGENITAL SINUS. EXPRESSION INCREASES WITH
RP GESTATIONAL AGE IN KIDNEY AND DUODENUM, BECOMING MAXIMAL IN
RP ADULTHOOD.
RN 19
RP P-TM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
RP AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN

```

Query Match	Best Local Similarity	Score 1262;	DB 1;	Length 411;
Matches 264;	Conservative 46;	Mismatches 86;	Indels 68;	Gaps 10;
7	CLLVIVSLLVCGGLACGPRGFG-KRRHKKTLPLAYKQFIPVNAKTLGASRYGSK	65		
13	CLFLILL--LLVPARGCGPGRVVGSRRRPRKPLVLAIKOPSPVPERKTLGASRYGSK	70		
66	ISRSSEFKELTPVNPDIIFKDEENTGADRLMTORCDKLNALASVNMOPGKTLVNT	125		
71	IARSSERKELTPVNPDIIFKDEENTGADRLMTORCDKLNALASVNMOPGKTLVNT	130		
126	EGMDEGHHSEESIHYGBRAVDITTSDRNSKTYGLARLAVAGEFDWVYYSKAHICSV	185		
131	EGMDEGHHSEESIHYGBRAVDITTSDRNSKTYGLARLAVAGEFDWVYYSKAHICSV	190		
186	KAESVVAKSGCGPGGATVHLBOGGTKLVNDSLGDGVLAADOGRLTYSDFLFLRD	245		
191	KSESAAKKTYGCGPPAGAOVRLNGERALASAVRGDVLNMGEGCTTFPSVLLFLDRE	250		
246	DGAKVEVYVETREPRERLLTAAHLTFVAPHNDSATREPAASGSGPPSGALPRALF	305		

Db 251 PNRLEAFVYETODPPRRLLATPAHLLFTA--DNHT-EPAHF-----RATF 294
 Qy 306 ASHVRPGORYVVAERDGRRLPAAVHSTVLTSEEAAGAPLTAOCTILINRLASCYA 365
 Db 295 ASHVQGOYVYV----SGVPGIQPARVAAS--THVALGSTAPLTRIGTLVEEDVVASCA 349
 Qy 366 VIEHSHMAHRAFPFRLAHALL-ALALAPARTDRGSGSGGGGGGAAVATPAGAADAP 424
 Db 350 AVADHLLAQLAEFWPLRFLPSLAWGSWTPSE----- 379
 Qy 425 GAGATAGIHWSQLYQIGTWLDSALHPICMAVKSXSRGAG 468
 Db 380 -----GVHWPQMLYRLGRLLEESTFHPLGMS-----GAG 410
 RESULT 12
 ID IHH_XENLA STANDARD; PRT; 409 AA.
 AC 091612;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Indian hedgehog protein precursor (IHH) (Banded hedgehog protein) (X-BHH).
 GN IHH OR BHH.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=95401852; PubMed=7671800;
 RA Ecker S.C., McGrew L.L., Lal C.-J., Lee J.J., von Kessler D.P.,
 RA Moon R.T., Beachy P.A.;
 RT "Distinct expression and shared activities of members of the hedgehog gene family of Xenopus laevis";
 RL Development 121:2337-2347(1995).
 CC -1- FUNCTION: SIGNAL INVOLVED IN THE EARLY INDUCTION AND PATTERNING OF ANTERODORSAL ECTODERM, NERVOUS SYSTEM AND SOMITES. INDUCES ECTOPIC CEMENT GLAND FORMATION IN EMBRYOS.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. HEDGEHOG PROTEIN IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- DEVELOPMENTAL STAGE: DETECTABLE WITHIN THE EARLY GASTRULA, AT STAGE 14 NEURULA, HIGH EXPRESSION IN NEUROECTODERM. EXPRESSED THROUGHOUT THE NEURAL PLATE AND SUBSEQUENTLY IN BOTH THE NERVOUS SYSTEM AND IN THE DERMATOME OF SOMITES.
 CC -1- INDUCTION: BY ACTIVIN.
 CC -1- PPM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPHOREOLYSIS ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U26404; AAA85165.1; -.
 DR HSSP: Q62226; IHH.

DR MEROPS; C46.0PW; -.
 DR Interpro: IPR000320; HH_signal.
 DR Interpro: IPR001767; Hedgehog_hinc.
 DR Interpro: IPR003586; Hedgehog_hinct.
 DR Interpro: IPR003587; Hedgehog_hinctn.
 DR Interpro: IPR001657; Sonichh.
 DR Pfam: PF01079; Hinc; 1.
 DR Pfam: PF01085; HH_signal; 1.
 DR PRINTS: PR00632; SONICHHOG.
 DR ProDom: PD003042; HH_signal; 1.
 DR SMART: SM00305; Hinc; 1.
 DR SMART: SM00306; Hinc; 1.
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 23
 FT CHAIN 24 409
 FT CHAIN 24 197
 FT CHAIN 198 409
 FT SITE 197 198
 FT SITE 267 267
 FT ACT_SITE 270 270
 FT BINDING 197 197
 FT DOMAIN 32 35
 FT LIPID 24 24
 SQ SEQUENCE 409 AA; 4591 MW; 6FB265367FB98627 CRC64;
 Query Match 50.5%; Score 1248; DB 1; Length 409;
 Best Local Similarity 56.3%; Pred. No. 7.1e-85;
 Matches 256; Conservative 53; Mismatches 92; Indels 54; Gaps 7;
 Qy 4 LARCLLVIVSSLLVSGSL-ACGPGKRGKRRPKTKTPLAYQFTIPNVAETLGSAGY 62
 Db 3 LPRVLLLCALALLLGAIVAGCCGPGKRVGRRRTKLSPLSTYQSPNVPEKTLGASGRY 62
 Qy 63 EGRISNRSEFKELTPYNDIIFKDEENTGADRLMTORCKRLNALAISVMQMPGVKL 122
 Db 63 EGRISNRSEFKELTPYNDIIFKDEITGADRLMTORCKRLNALAISVMQMPGVKL 122
 Qy 123 RYEGWDEDHHEESLHYGKRAVDITTSRDSKTKGMLARLAVEAGFDWYVESNAHIH 182
 Db 123 RYEGWDEDHHEESLHYGKRAVDITTSRDRNKKYGMRLAVEAGFDWYVESNAHIH 182
 Qy 183 CSYKAENSVAKGCGPGSGATVHLBOGGTKLVKDISPGRVLAADQGRLLVSDFTFL 242
 Db 183 CSYKAENSVAKGCGPGSGATVHLBOGGTKLVKDISPGRVLAADQGRLLVSDFTFL 242
 Qy 183 CSYKSEHSAAKGCGCPGALATLSEGEKIPQSLSFGRLVLAANDSGRPTSDLSFL 242
 Db 243 DRDDGAKYFYVITREPRERLLLTAAHLLFVAPHNDSATGEPEASSGSPSGALGPR 302
 Qy 243 DHSPEKHEHMPVKTQDPHRRFLFTRAHLLFVSDNVTSPASERO----- 286
 Db 303 ALASRYRPGORYVVAERDGRRLPAAVHSTVLTSEEAAGAPLTAOCTILINRLAS 362
 Qy 287 AVPAASVPRQYI-LVSNVG---LIPAKVSVN-TQTNGAAPLPHOHTLVVDVVS 341
 Db 363 CVAVIEESHMAHRAFPFRLAHALLALAPARTDRGSGSGGGGGGAAVATPAGAAD 422
 Qy 342 CFAVYQKQRLAQLVWPLRLVNL-----GIATGP----- 373
 Db 423 APGAGATAGIHWSQLYQIGTWLDSALHPICMA 457
 Db 374 -----SOOMGIHWSKALYHLGRLILGNENHPIGI 404
 RESULT 13
 ID IHH_BRARE STANDARD; PRT; 412 AA.
 AC 098862;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Indian hedgehog protein precursor (IHH) (Echidna hedgehog protein).

GN IHH OR EHH.
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 ON NCBI_Taxid=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96310864; PubMed=8684485;
 RA Currie P.D., Ingham P.W.;
 RT "Induction of a specific muscle cell type by a hedgehog-like protein in zebrafish."
 RL Nature 382:452-455(1996).
 RN [2]
 RP SEQUENCE OF 113-170 FROM N.A.
 RC TISSUE=Muscle;
 RX MEDLINE=97075114; PubMed=8917540;
 RA Zardoya R., Abouheif E., Meyer A.;
 RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species closely related to the zebrafish."
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF NOTOCHORD EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE NOTOCHORD THAT INDUCES SOMITE PATTERNING AND MUSCLE PIONEER DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM.
 CC -1- TISSUE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT THE MID-GASTRULA STAGE. DISAPPEARS AT THE END OF THE SOMITOGENESIS.
 CC -1- PM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to license@isb-sdb.ch).
 CC -----
 DR EMBL; Y08426; CAA69702.1; .
 DR HSSP; O62226; 1VHN.
 DR MEROPS; C46; UPW; .
 DR ZFIN; ZDB-GENE-980526-135; ehb.
 DR InterPro; IPR000320; HH_signal.
 DR InterPro; IPR001767; Hedgehog_hntc.
 DR InterPro; IPR003586; Hedgehog_hntc.
 DR InterPro; IPR003587; Hedgehog_hntn.
 DR InterPro; IPR002203; Intein.
 DR InterPro; IPR001657; Sonichh.
 DR Pfam; PF01079; Hntc; 1.
 DR Pfam; PF01083; HH_signal; 1.
 DR PRINTS; PR006342; SONICHOG.
 DR ProDom; PD003042; HH_signal; 1.
 DR SMART; SM00305; Hntc; 1.
 DR SMART; SM00306; Hntn; 1.
 DR PROSITE; PS50817; INTEIN_N_TER; 1.
 KM Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 412 INDIAN HEDGEHOG PROTEIN.
 FT CHAIN 24 197 INDIAN HEDGEHOG PROTEIN N-PRODUCT.

FT CHAIN 198 412 INDIAN HEDGEHOG PROTEIN C-PRODUCT.
 FT SITE 197 198 CLEAVAGE (AUTO-) (BY SIMILARITY).
 FT SITE 246 246 INVOLVED IN CHOLESTEROL TRANSFER (BY SIMILARITY).
 FT SITE 269 269 INVOLVED IN AUTO-CLEAVAGE (BY SIMILARITY).
 FT ACT_SITE 272 272 ESSENTIAL FOR AUTO-CLEAVAGE (BY SIMILARITY).
 FT BINDING 197 197 CHOLESTEROL (BY SIMILARITY).
 FT LIPID 24 24 PALMITATE (BY SIMILARITY).
 FT CONFLICT 121 121 R -> K (IN REF. 2).
 SQ SEQUENCE 412 AA; 45746 MW; 21D43F05A278CA1 CRC64;
 Query Match 49.1%; Score 1213; DB 1; Length 412;
 Best Local Similarity 57.4%; Pred. No. 2,76-82;
 Matches 251; Conservative 44; Mismatches 86; Indels 56; Gaps 7;
 QY 24 CGPGRGFKRRHPKLTPLATKQIPNVAEITLGAASGREGKISNSREFELTPNYPD 83
 DB 24 CGPGRGFKRRHPKLTPLATKQIPNVAEITLGAASGREGKISNSREFELTPNYPD 83
 QY 84 IIFDEENTGADRLMTORCKDKLNALATSVNOMPVGKLTATGEGDEGHSESLHYEG 143
 DB 84 IIFDEENTGADRLMTORCKDKLNALATSVNOMPVGKLTATGEGDEGHSESLHYEG 143
 QY 144 RAVDITTSRDRSKYGMRLAVEAGFDVYVESKAHICGVKANSYAAKSGCFFPSA 203
 DB 144 RAVDITTSRDRSKYGMRLAVEAGFDVYVESKAHICGVKANSYAAKSGCFFPSA 203
 QY 204 TVHEOGSTKLVKDLSPEDRYLA---ADDOGRLLYSDFLTLDKDGAKKVFYIETREP 260
 DB 204 TVHEOGSTKLVKDLSPEDRYLA---ADDOGRLLYSDFLTLDKDGAKKVFYIETREP 260
 QY 261 RERLITTAHLLEFVAHPNDSATGPEASSGSGPSCGLPRALEFASVPRGORYVAE 320
 DB 261 RERLITTAHLLEFVAHPNDSATGPEASSGSGPSCGLPRALEFASVPRGORYVAE 320
 QY 321 RDGRRLLPAVHVSTLSEENAGAVAPLTAGTILINVLASCAVIEHSHMARAPF 380
 DB 321 RDGRRLLPAVHVSTLSEENAGAVAPLTAGTILINVLASCAVIEHSHMARAPF 380
 QY 309 --GKLRKPSQITVHGVNED--GGLVPLPLANGTVVNDVLSCAAVNRQLAHMAFAPL 365
 DB 309 --GKLRKPSQITVHGVNED--GGLVPLPLANGTVVNDVLSCAAVNRQLAHMAFAPL 365
 QY 381 RLARLLAALPARTRDRGDSGGGDRGGGGRVALTPAGAADAPCAAGTACIHWTSQLY 440
 DB 381 RLARLLAALPARTRDRGDSGGGDRGGGGRVALTPAGAADAPCAAGTACIHWTSQLY 440
 QY 441 QIGTWLDPSEALPLGM 457
 DB 441 QIGTWLDPSEALPLGM 457
 QY 391 GIGKLLIDSELFHPLAL 407
 DB 391 GIGKLLIDSELFHPLAL 407
 RESULT 14
 DHN_MOUSE STANDARD; PRT; 396 AA.
 AC 061488;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Desert hedgehog protein precursor (DHN) (HMG-3).
 GN DHN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv;
 RX MEDLINE=94094334; PubMed=7916661;
 RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J., McMahon J.A., McMahon A.P.;
 RT "Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity."
 RL Cell 75:1417-1430(1993).
 RN [2]

SEQUENCE OF 120-168 FROM N.A.
 MEDLINE-9536997; PubMed-7720571;
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Smandl B.K.,
 RA Zhao R., Seidman M.F., Fallon J.F., Beachy P.A.;
 RT "Products, genetic linkage and limb patterning activity of a murine
 hedgehog gene";
 RL Development 120:3339-3353(1994).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A
 CC SPERMATOCYTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES
 CC DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ADULT TESTES. NOT EXPRESSED IN
 CC LIMB BUDS.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC in the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----
 CC EMBL: X76292; CAA53924.1; -;
 CC DR HSSP: Q62226; LVNH.
 CC DR MEROPS: C46.004; -;
 CC DR MGD: MGI:94891; Dhh.
 CC DR InterPro: IPR000320; HH_signal.
 CC DR InterPro: IPR001767; Hedgehog_hhnc.
 CC DR InterPro: IPR003586; Hedgehog_hhnc.
 CC DR InterPro: IPR003587; Hedgehog_hhnc.
 CC DR InterPro: IPR002203; Intein.
 CC DR InterPro: IPR001657; Sonichh.
 CC DR Pfam: PF01079; Hhnt; 1.
 CC DR Pfam: PF01085; HH_signal; 1.
 CC DR PRINTS: PRO0632; SONICHHOG.
 CC DR ProDom: PD003042; HH_signal; 1.
 CC DR SMART: SM00305; HhntC; 1.
 CC DR SMART: SM00306; HhntN; 1.
 CC DR PROSITE: PS00817; INTEIN_NTER; 1.
 CC DR Developmental protein: Autocatalytic cleavage; Hydrolyase; Protease;
 CC Signal; Lipoprotein; Palmitate.
 CC KW
 CC FT SIGNAL 1 22
 CC FT CHAIN 23 396
 CC FT CHAIN 23 198
 CC FT CHAIN 199 396
 CC FT SITE 198 199
 CC FT SITE 244 244
 CC FT SITE 268 268
 CC FT SITE 271 271
 CC FT ACT_SITE 271 271
 CC FT BINDING 198 198
 CC FT LIPID 23 23
 CC FT SEQUENCE 396 AA; 43542 MW; AFEFB051BE950FD8 CRC64;
 Query Match 47.9%; Score 1183; DB 1; Length 396;
 Best Local Similarity 54.5%; Pred. No. 4; 3e-80;
 Matches 244; Conservative 55; Mismatches 87; Indels 62; Gaps 10;

QY 3 LIAELLVAVSSLLVSCGLACGPGRG-FGKRHRPK-LTPYAKQFIPNVAEKTGASG 60
 DB 7 LIPICL-----ALLASQSCGPGRGVRRYRKQLVPLLYKQFVSMBERTLGASG 61
 QY 61 RYEGKSRNSRSEKELTPYNDIIFKDEENTGADRLMTQCKDKINALAISVMNMPGV 120
 DB 62 PAGRVTRESEFRDLVPYNDIIFKDEENSGADRLMERCKERYNALAIIVMMNMPGV 121
 QY 121 KLRVTEGMEDDHHSESEHYEGRAVDITSPDRSKYGLRLVAEAGFDWYTESKAH 180
 DB 122 RLRYTEGMEDDHHADSDHYEGRALDITSPDRSKYGLRLVAEAGFDWYTESKAH 181
 QY 181 HCSVAENSVAASGCGCPGSAVTHLEOGGKRLVNDLSPGDRVLAADOGRLYSDFLT 240
 DB 182 HVSVAENSLVAVRAGCGCPGSAVTHLEOGGKRLVNDLSPGDRVLAADOGRLYSDFLT 241
 QY 241 FLDRDGAARVYVETREPRERLITLAHLFVAPNDSATGEPAASGSGP-PGCGAL 299
 DB 242 FLDRDGAARVYVETREPRERLITLAHLFVAPNDSATGEPAASGSGP-PGCGAL 299
 QY 300 GPRLPASPVRPGQRYVVAERDGRRLPAVHVSYLEEAGAVAPLTAAGTILINRY 359
 DB 287 AP--VFARLRAGDSVLA---PGDALOPARVARYA-REAVGVAPLTAAGTILINRY 339
 QY 360 LASCVAIEHSVNRAPFAPFLAHLAALAPARTDRCGSGCGGCGGVALTAGP 419
 DB 340 LASCVAIEHSVNRAPFAPFLAHLAALAPARTDRCGSGCGGCGGVALTAGP 419
 QY 420 AADAPGAGTAGIHWYSQLYQIGTWL 447
 DB 373 GAVQP-----TGMHWYSRLYLAEELM 395
 RESULT 15
 DHH_HUMAN STANDARD; PRT; 396 AA.
 AC 043323; Q15794;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Desert hedgehog protein precursor (Dhh) (HhG-3).
 GN DHH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RA "Sequence from N.A.
 RA Tate G., Kishimoto K., Mitsuura T.;
 RT "Expression of Sonic hedgehog and its receptor patched/smoothed in
 RT human cancer cell lines and embryonic organs";
 RL J. Biochem. Mol. Biol. Biophys. 4:27-34(2000).
 RN [2]
 RA "Sequence of 85-178 from N.A.
 RA Tissue-Kidney;
 RA Drummond I.A.;
 RT "Human desert hedgehog";
 RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RA DISEASE;
 RA PubMed-11017805;
 RA Uehara F., Tate G., Itoh K., Yamaguchi N., Douchi T., Mitsuura T.,
 RA Osame M.;
 RT "A novel mutation of desert hedgehog in a patient with 46,XY partial
 RT gonadal dysgenesis is accompanied by multifascicular neuropathy";
 RL Am. J. Hum. Genet. 67:1302-1305(2000).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A
 CC SPERMATOCYTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES
 CC DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:06:28 ; Search time 30.5535 Seconds
(without alignments)
3203.314 Million cell updates/sec

Title: US-09-827-110A-15
Perfect score: 2469
Sequence: 1 MLLARCLLVSLVCS.....GMAYKSKRGAGGAREGA 475

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1544.5	62.6	434	13	057404 pleurodeles
2	1454	58.9	414	13	09W709 pleurodeles
3	1308	53.0	406	13	057567 pleurodeles
4	1094.5	44.3	415	5	09U526 branchiostoma
5	1084.5	43.9	415	5	017499 branchiostoma
6	1049.5	42.2	442	13	073803 fugu rubrip
7	1041.5	42.2	447	5	09G8A8 gryllus blm
8	916	37.1	410	5	061676 lytechinus
9	850	34.4	161	11	09R179 ratius norv
10	733	29.7	150	11	09Y6V7 ambystoma m
11	731	29.6	139	6	09X5I6 bos taurus
12	709	28.7	138	11	09W6C1 eleutheroda
13	702	28.4	177	11	09WV29 ratius norv
14	619	25.1	185	5	096699 junonia coe
15	531	21.5	99	13	08QGN4 scylliorhinu
16	523	21.2	119	13	042128 oryzias lat

17	511	20.7	129	11	09WUP6 ratius norv
18	476.5	19.3	137	13	042234 coturnix co
19	392	15.9	80	13	042441 oryzias lat
20	392	15.9	87	5	09TX30 anopheles g
21	339	13.7	63	13	091814 rana catesb
22	332	13.4	64	13	09PRF5 oryzias lat
23	288	12.1	60	5	0952D8 hydra atten
24	289	11.7	56	13	0902R0 oryzias lat
25	237	9.6	49	5	09TX33 hirtudo medl
26	229	9.3	45	5	09TX31 tribolium c
27	188	7.6	54	13	042233 coturnix co
28	187.5	7.6	125	11	09ESH3 ratius norv
29	175.5	7.1	48	5	09TX32 stronglylloe
30	163.5	6.6	557	5	094410 caenorhabd1
31	163.5	6.6	557	5	094129 caenorhabd1
32	159.5	6.5	615	5	P91573 caenorhabd1
33	149	6.0	1207	5	021535 caenorhabd1
34	138.5	5.6	790	5	022872 caenorhabd1
35	138.5	5.6	1203	5	021835 caenorhabd1
36	135.5	5.5	485	5	094128 caenorhabd1
37	134.5	5.4	481	5	045992 caenorhabd1
38	128	5.2	550	5	045273 caenorhabd1
39	126.5	5.1	1021	5	09XUV2 caenorhabd1
40	123.5	5.0	550	5	094130 caenorhabd1
41	123.5	5.0	2639	5	076786 antherea p
42	121	4.9	868	5	09XV14 caenorhabd1
43	119	4.8	318	16	086703 streptomyce
44	119	4.8	2655	5	0964F4 antherea y
45	118	4.8	1832	3	08TGH8 podospora a

ALIGNMENTS

RESULT 1
ID 057404 PRELIMINARY; PRT; 434 AA.
AC 057404;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Sonic hedgehog-related protein.
GN TW-SHH.
OS Pleurodeles waltlil (Iberian ribbed newt).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroideae; Salamandridae;
OC Pleurodeles.
OX NCBI_TaxID=8319;
RN [1]
RP SEQUENCE FROM N.A.
RA Caubit X., Nicolas S., Le Parco Y.;
RT "Pleurodeles sonic hedgehog";
RT Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003532; AAB94412.1; -.
DR HSSP; Q62226; 1VHR.
DR MEROPS; C46.002; -.
DR InterPro: IPR001767; Hedgehog_hnt.
DR InterPro: IPR003586; Hedgehog_hntc.
DR InterPro: IPR003587; Hedgehog_hntw.
DR InterPro: IPR003320; Hh_signal.
DR InterPro: IPR002203; Intein.
DR InterPro: IPR001657; SonichH.
DR Pfam; PF01085; Hh_signal; 1.
DR Pfam; PF01079; Hnt; 1.
DR PRINTS; PRO0632; SONICHOG.
DR PRODOM; PD003042; Hh_signal; 1.
DR SMART; SM00305; Hntc; 1.
DR SMART; SM00306; Hntw; 1.
SQ SEQUENCE 434 AA; 48421 MW; A9495E367151AE74 CRC64;

Query Match 62.6%; Score 1544.5; DB 13; Length 434;
Best Local Similarity 67.7%; Pred. No. 8.1e-108;
Matches 313; Conservative 41; Mismatches 77; Indels 31; Gaps 6;

```

OY 1 MLLARCLLLVYSSLLVCSGLACGPRGGRKRRHPKPLPYPVAKETTLGASG 60
Db 4 MILLRRDLGGFISTLLVPLGLCGCGPRGIGORRRPKPLPLAKOTIPVSEKTLGASG 63
OY 61 RYBKISRNSERFEKELTPNPNPDIIFKDEENTGADRLMTCORKDLNALAI SVANOMPVG 120
Db 64 RYEVKATITNSERFEPOLTPNPNPDIIFKVEEETGADRLMTCORKDLNLSI SVANOMPVG 123
OY 121 KLRTGEGDEGHHSEESLHNEGAVDITSDRDRSKYGMILARLAVAGDWWYFEKAK 180
Db 124 KLRTGEGDEGHHSEESLHNEGAVDITSDRDRSKYGMILARLAVAGDWWYFEKAK 183
OY 181 IHCYKAKENSVAAKSGCGFPGSARVHLEOGGTLYKPLSGDRVLADDDGRLYSDFLT 240
Db 184 IHCYKAKENSVAAKSGCGFPGSARVHLEOGGTLYKPLSGDRVLADDDGRLYSDFLT 243
OY 241 FLDDDDGAKKYFYIETREPERRLTLTAHLLEVPAPINDSNTGEPKSSGSGPPSGALG 300
Db 244 FMDEEETARKFYVIETSLPERDRRLTLTAHLLEFVAOEHPC-----NAAENFRSKFGR 298
OY 301 PRALEFASVRRCGVVVAERDDGRRLTLPAAVHSVTLSSEAGAYALTLNOSTLLNRYL 360
Db 299 FRSMPASSVRGHHV-LTEDEEG-RGIREKLTVDRLYL-EETGTATAVYLAHGTVIDRVL 355
OY 361 ASCYAVIIEHSHMAHRAFPRLAHLAALLAALAPARTDRGSDSGGDRGGGGRYALTAPGA 420
Db 356 ASCYAVIIEHSHMAHRAFPRLVCGILSPSP--QDYSMSH----- 394
OY 421 ADAPGAGTATGHHYSOLLQVQIGWLLDSEALHPLGAAVSS 462
Db 395 --PAPSOABSVHHYSELTIRIGTWVLAQDTIHLPLGAAVSS 434

```

RESULT 2				
09W709	ID	09W709	PRELIMINARY;	PRT; 414 AA.
AC	09W709;	01-NOV-1999 (TRMBLrel. 12, Created)		
DT	01-NOV-1999 (TRMBLrel. 12, Last sequence update)			
DT	01-MAR-2002 (TRMBLrel. 20, Last annotation update)			
DE	Sonic hedgehog.			
GN	SHH.			
OS	Paralichthys olivaceus (Flounder).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC	Acanthomorphi; Acanthopterygii; Percomorphi; Pleuronectiformes;			
OC	Pleuronectoidei; Paralichthyidae; Paralichthys.			
OX	NCBI_TaxID=8255;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE:99238226; PubMed-10223710;			
RA	Suzuki T., Ichiro O., Kurokawa T.;			
RT	"Retinoic acid given at late embryonic stage depresses sonic hedgehog			
RT	and Hoxd-4 expression in the pharyngeal area and induces skeletal			
RT	malformation in flounder (Paralichthys olivaceus) embryos."			
RL	Dev. Growth Differ. 41:143-152(1999).			
DR	EMBL: AB029748; BAA82360.1; -			
DR	HSPS: Q62226; IYHH.			
DR	MEROPS: C46.001; -			
DR	InterPro: IPR001767; Hedgehog_hnt.			
DR	InterPro: IPR003586; Hedgehog_hntc.			
DR	InterPro: IPR003587; Hedgehog_hlntN.			
DR	InterPro: IPR000320; HH_signal.			
DR	InterPro: IPR002203; Intein.			
DR	InterPro: IPR001657; SonicHH.			
DR	Pfam: PF01085; HH_signal; 1.			
DR	Pfam: PF01079; HlntL; 1.			
DR	PRINTS: PR006332; SONICHOG.			
DR	ProDom: PD003042; HH_signal; 1.			
DR	SMART: SM00305; HlntC; 1.			
DR	SMART: SM00306; HlntN; 1.			
SO	SEQUENCE 414 AA; 45945 MW; 50607BF3DB7CDDA3 CRC64;			

Query Match	58.9%	Score 1454;	DB 13;	Length 414;
Best Local Similarity	62.4%	Pred. NO. 4.7e-101;		
Matches 289; Conservative	55;	Mismatches 69;	Indels 50;	Gaps 77;

[illegible]

RESULT 3	ID	057567	PRELIMINARY;	PRT;	406 AA.
AC	057567;				
DT	01-JUN-1998 (TREMBLrel. 06, Created)				
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)				
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)				
DE	Hedgehog segment polarity homolog.				
OS	Notophthalmus viridescens (Eastern newt). (Triturus viridescens).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Amphibia; Batrachia; Caudata; Salamandroidae; Salamandridae;				
OC	Notophthalmus				
OX	NCBI_TaxID=8316;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Stark D.R., Gates P.B., Brookes J.P., Ferretti P.;				
RT	"Hedgehog homologue from Notophthalmus viridescens.";				
RL	Dev. Dyn. 0:0-0(1998)				
CC	-1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE				
CC	PHOSPHORIBOSYLTRANSFERASE FAMILY.				
DR	EMBL; AF047466; AAC03108.1; -.				
DR	HSSP; AE6226; 1VHH.				
DR	InterPro; IPR001767; Hedgehog_hint.				
DR	InterPro; IPR003586; Hedgehog_hintC.				
DR	InterPro; IPR003587; Hedgehog_hintN.				
DR	InterPro; IPR000320; HH_signal.				
DR	InterPro; IPR002375; pr/py_rp_transf.				
DR	InterPro; IPR001657; SonicHH.				
DR	Pfam; PF01085; HH_signal; 1.				
DR	Pfam; PF01079; Hint; 1.				
DR	PRINTS; PRO0632; SONICHHOG.				
DR	ProDom; PD003042; HH_signal; 1.				
DR	SMART; SM00305; HintC; 1.				
DR	SMART; SM00306; HintN; 1.				
DR	PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.				

KW Glycosyltransferase; Transferase.
SQ SEQUENCE 406 AA; 45072 MW; 5842CCAD5314D4D0 CRC64;

Query Match 53.0%; Score 1308; DB 13; Length 406;
Best Local Similarity 59.2%; Pred. No. 4.2e-90;

Matches 274; Conservative 41; Mismatches 80; Indels 66; Gaps 9;

```

QY 1 MLLARCLLVVSSLLVCSG--LACGPRGFGKRRRHKLTPLAYKQFIPNVAEKTGLAS 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 MKLLAPPLAVNACALLGGGALGCGPRGIRGRPRRLPIPLSTYQFLPHVEKTLGAS 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 60 GREGKISNSERFKELTPYNDITFKDEENTGADRLMTORCKDKLNLAISVNMQWG 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 GREGKISNSERFKELTPYNDITFKDEENTGADRLMTORCKDKLNLAISVNMQWG 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 120 VKLRVTEGDEDOGHSEESLHTEGRAVDITTSRDORSKGMRLARLVEAGFDWVYESKA 179
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 121 VKLRVTEGDEDOGHSEESLHTEGRAVDITTSRDORSKGMRLARLVEAGFDWVYESKA 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 180 HHCYSVKAENSVAASKGCGPFGSATVHLBOGGTKLVKDLSPGDRVLAADOGRLYSDFL 239
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 181 HHCYSVKAENSVAASKGCGPFGSATVHLBOGGTKLVKDLSPGDRVLAADOGRLYSDFL 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 240 TFLDRDGAKKVYVYETREPRERLLTAHLLFVA-----PHNDSATGEPPEASGSGGPP 294
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 241 TFLDRDSTAKKEVYVETRDPRRLALTAHLLFVADNFTVPLTDFST----- 288
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 295 SGGALPRLAFASRVPRGQVYVVAERDGRLLPAVHSTVLSSEAGAVAPLTAOGTI 354
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 289 -----VFASHVQPGQ--YILTE--GYLGQIPARVSVT--TQIDSGAYALTSIGTL 334
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 355 LNRVLASCYAVIEHSMAHRAFPARLALHALALAPARTDRGDSGGGDRGGGGGRVA 414
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 335 LVDDVYVSCFAYVQKHQALQALFMPRLTLYHSV-----GRPE 370
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 415 LTRPGAADAPGAGATAGIHWYSOLLYQIGTWLIDSEALHPLGM 457
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 371 -TQP-----EGMHVSLILYRLGKVLMLKQFHPFGM 401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 4

Q90526 PRELIMINARY; PRT; 415 AA.

```

AC Q90526;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hedgehog protein.
GN Hedgehog.
OS Branchiostoma belcheri (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_Taxid=7741;
RN [1]
RP SEQUENCE FROM N.A.
RA Mao B., Zhang Y., Zhang S., Zhang H.;
RT "A preliminary study on qingdao amphioxus hedgehog gene.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ245882; CAB58231.2; -
DR HSSP; Q62226; 1VHH.
DR InterPro; IPR001767; Hedgehog_hint.
DR InterPro; IPR003586; Hedgehog_hintc.
DR InterPro; IPR003587; Hedgehog_hintn.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR002203; Intein.
DR InterPro; IPR001657; SonICHH.
DR Pfam; PF01085; HH_signal; 2.
DR PRINTS; PR00632; SONICHHOG.
DR PRODOM; PD003042; HH_signal; 1.
DR SMART; SM00305; Hintc; 2.
DR SMART; SM00306; Hintn; 2.
SQ SEQUENCE 415 AA; 46912 MW; BF864704752BD6A8 CRC64;

```

Query Match 44.3%; Score 1094.5; DB 5; Length 415;
Best Local Similarity 50.2%; Pred. No. 4.7e-74;
Matches 231; Conservative 57; Mismatches 111; Indels 61; Gaps 10;

```

QY 3 LLARCLL-LVYSSLLVCSG--LACGPRGFGKRRRHKLTPLAYKQFIPNVAEKTGLAS 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4 VLARWVTVLVALTALCTHNGPSEACGPRGIRGRPRRLTFLPVYKQMPAVSENFEGAS 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 60 GREGKISNSERFKELTPYNDITFKDEENTGADRLMTORCKDKLNLAISVNMQWG 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 64 GLENGRITDSEKERTLKQNFNDITFKDEENTGADRLMTORCKDKLNLAISVNMQWG 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 120 VKLRVTEGDEDOGHSEESLHTEGRAVDITTSRDORSKGMRLARLVEAGFDWVYESKA 179
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 124 VKLRVTEGDEDOGHSEESLHTEGRAVDITTSRDORSKGMRLARLVEAGFDWVYESKA 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 180 HHCYSVKAENSVAASKGCGPFGSATVHLBOGGTKLVKDLSPGDRVLAADOGRLYSDFL 239
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 184 HHCYSVKAENSVAASKGCGPFGSATVHLBOGGTKLVKDLSPGDRVLAADOGRLYSDFL 243
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 240 TFLDRDGAKKVYVYETREPRERLLTAHLLFVA-PHNDSATGEPPEASGSGPFGAL 299
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 244 TFLDRDGRGPRWYVYETREPRERLLTAHLLFVA-PHNDSATGEPPEASGSGPFGAL 285
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 300 GPR-ALFASRVPRGQVYVVAERDGR--RRLPAVHSTVLSSEAGAVAPLTAOGTI 356
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 286 APRVAFMSDVRGGE--FLTPESDGGGFRKY---ELVSTMRRE-KGAAVAPLVHGTIVV 340
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 357 NRVLASCYAVIEHSMAHRAFPARLALHALALAPARTDRGDSGGGDRGGGGGRVALT 416
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 341 DNYAMSCYALIEQALAHVFAFPRLYYQILTSSL----- 374
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 417 APGAADAPGAGAT--AGIHWYSOLLYQIGTWLIDSEALHPL 454
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 375 ----WDGPHSDOTLQEGVHWYPSFFYRYGISLVEPTLLHP 410
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 5

O17499 PRELIMINARY; PRT; 415 AA.

```

AC O17499;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Amphihh protein.
GN AMPHIHH.
OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_Taxid=7739;
RN [1]
RP SEQUENCE FROM N.A.
RA Shmied S.M.;
RT "Sequence and expression of an amphioxus hedgehog gene.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y13858; CAA74169.1; -
DR HSSP; Q62226; 1VHH.
DR InterPro; IPR001767; Hedgehog_hint.
DR InterPro; IPR003586; Hedgehog_hintc.
DR InterPro; IPR003587; Hedgehog_hintn.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR002203; Intein.
DR InterPro; IPR001657; SonICHH.
DR Pfam; PF01085; HH_signal; 1.
DR Pfam; PF01079; Hintc; 1.
DR PRINTS; PR00632; SONICHHOG.
DR PRODOM; PD003042; HH_signal; 1.
DR SMART; SM00305; Hintc; 1.
DR SMART; SM00306; Hintn; 1.
SQ SEQUENCE 415 AA; 46765 MW; 5EBD2BD2158DD931 CRC64;

```

Query Match 43.9%; Score 1084.5; DB 5; Length 415;


```
|||||
Db 1 IIFKEENTGADRLMTQCKDKLNALAI SYNQWGVKRLRYEGDEGHHSEESLHYEG 60
Oy 144 RAYDITTSDDRBSKYGMLARLAVEAGFDWVYYESKAHHCYSKAKNSVAASKGCGCPGSA 203
Db 61 RAYDITTSDDRBSKYGMLARLAVEAGFDWVYYESKAHHCYSKAKNSVAASKGCGCPGSA 120
Oy 204 TVHLEGGCTKLVKDLSPGDRVLAADQGRLLYSDFLTFELDR 244
Db 121 TVHLEGGCTKLVKDLSPGDRVLAADQGRLLYSDFLTFELDR 161

RESULT 10
Oy 09YGV7 PRELIMINARY; PRT; 150 AA.
AC 09YGV7;
DT 01-MAY-1999 (TREMBLREL. 10, Created)
DT 01-MAY-1999 (TREMBLREL. 10, Last sequence update)
DT 01-JUN-2002 (TREMBLREL. 21, Last annotation update)
DE Sonic hedgehog (Fragment).
OS Ambystoma mexicanum (Axolotl).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidae; Ambystomatidae;
OC Ambystoma
OX NCBI_TaxID=8296;
RN [1]
RP SEQUENCE FROM N.A.
RA Torok M.A., Izpizua-Belmonte J.C., Gardiner D.M., Bryant S.V.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF031480; AAD18128.1; -
DR HSSP; O62226; IVHH.
DR InterPro; IPR001167; Hedgehog_hntc.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; SonicHH.
DR Pfam; PF01085; HH_signal; 1.
DR Pfam; PF01079; Hntc; 1.
DR PRINTS; PR00632; SONICHHOG.
DR PRODOM; PD003042; HH_signal; 1.
FT NON_TER 1 150
SQ SEQUENCE 150 AA; 16599 MW; 986F65037A69A8AC CRC64;

Query Match
Best Local Similarity 93.38; Score 733; DB 13; Length 150;
Matches 140; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Oy 83 DIFDEENTGADRLMTQCKDKLNALAI SYNQWGVKRLRYEGDEGHHSEESLHYE 142
Db 1 DIFDEENTGADRLMTQCKDKLNALAI SYNQWGVKRLRYEGDEGHHSEESLHYE 60
Oy 143 GRAVDITTSDDRBSKYGMLARLAVEAGFDWVYYESKAHHCYSKAKNSVAASKGCGCPGSA 202
Db 61 GRAVDITTSDDRBSKYGMLARLAVEAGFDWVYYESKAHHCYSKAKNSVAASKGCGCPGSA 120
Oy 203 ATVHLEGGCTKLVKDLSPGDRVLAADQGR 232
Db 121 AKVLEHGVTRPVADLRGDRVLAADQGR 150

RESULT 11
Oy 09XSI6 PRELIMINARY; PRT; 139 AA.
AC 09XSI6;
DT 01-NOV-1999 (TREMBLREL. 12, Created)
DT 01-NOV-1999 (TREMBLREL. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLREL. 17, Last annotation update)
DE Sonic hedgehog (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC TISSUE=TOOTH GERM;
RA Koyama E., Iwamoto M., Ohmori T., Kurisu K., Wu C., Ookura T.,
RA Bashir M.M., Tucker T., Pacifici M.;
RT "Development of Stratum intermedium and its role as a Sonic Hedgehog-
RT Signaling Structure During Odonogenesis."
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF144100; AAD33926.1; -
DR HSSP; O62226; IVHH.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; SonicHH.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHHOG.
DR PRODOM; PD003042; HH_signal; 1.
FT NON_TER 1 139
SQ SEQUENCE 139 AA; 15961 MW; C83BDB9682046B1 CRC64;

Query Match
Best Local Similarity 98.68; Score 731; DB 6; Length 139;
Matches 137; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 43 AYKQITPVNAEKTLAGASRGYEGKISRNSERFKELTPNPNPDIIFFDEENTGADRLMTQRC 102
Db 1 AYKQITPVNAEKTLAGASRGYEGKISRNSERFKELTPNPNPDIIFFDEENTGADRLMTQRC 60
Oy 103 KDKLNALAI SYNQWGVKRLRYEGDEGHHSEESLHYEGRAVDITTSDDRBSKYGMLA 162
Db 61 KDKLNALAI SYNQWGVKRLRYEGDEGHHSEESLHYEGRAVDITTSDDRBSKYGMLA 120
Oy 163 RLAVEAGFDWVYYESKAHI 181
Db 121 RLAVEAGFDWVYYESKAHI 139

RESULT 12
Oy 09W6C1 PRELIMINARY; PRT; 138 AA.
AC 09W6C1;
DT 01-NOV-1999 (TREMBLREL. 12, Created)
DT 01-NOV-1999 (TREMBLREL. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLREL. 17, Last annotation update)
DE Sonic hedgehog protein (Fragment).
GN SHH.
OS Eleutherodactylus coqui (Puerto Rican coqui).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Leptodactylidae;
OC Eleutherodactylus.
OX NCBI_TaxID=57060;
RN [1]
RP SEQUENCE FROM N.A.
RA Carl T.F., Richardson M.K., Olsson L., Schlosser G., Klymkowsky M.W.,
RA Hanken J.;
RT "Differences in vertebrate limb development revealed by studies of the
RT direct developing frog E. coqui."
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF113403; AAD23436.1; -
DR HSSP; O62226; IVHH.
DR InterPro; IPR001657; HH_signal.
DR InterPro; IPR000320; HH_signal.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHHOG.
DR PRODOM; PD003042; HH_signal; 1.
FT NON_TER 1 138
SQ SEQUENCE 138 AA; 15751 MW; FF4156A17F4681F0 CRC64;

Query Match
Best Local Similarity 96.48; Score 709; DB 13; Length 138;
Matches 133; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 48 IPNVAEKTLAGASRGYEGKISRNSERFKELTPNPNPDIIFFDEENTGADRLMTQCKDKLN 107
```

```
DB 1 IPVNAKTLGASRGYEGKITRNSERKELTPNYSNIIIFKDEENTATADRLMORCKDKLN 60
QY 108 ALAISVNMNPGVKLVNTEGWDGHHSESLHYEGRAVDITTSDBRSKYGMARLAVE 167
DB 61 ALAISVNMNPGVKLVNTEGWDGHHSESLHYEGRAVDITTSDBRSKYGMARLAVE 120
QY 168 AGFDWYVESKAHICSV 185
DB 121 AGFDWYVESKAHICSV 138

RESULT 13
Q9WV29 PRELIMINARY: PRT: 177 AA.
ID 09WV29;
AC 09WV29;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Indian hedgehog protein (Fragment).
OS Rattus norvegicus (Rat).
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBITaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RA Garges P.L., Meyer R.A., Jr., Brown C.A., Price D.K.;
RT "Indian hedgehog in rat.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DDA databases.
DR EMBL: AF163914; AAD45372.1; -.
DR HSSP: 062226; 1VH.
DR MEROPS: C46.003; -.
DR InterPro: IPR001767; Hedgehog_hint.
DR InterPro: IPR003587; Hedgehog_hint.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR001657; SonicHH.
DR Pfam: PF01085; HH_signal; 1.
DR PRINTS: PR00632; SONICHHOG.
DR ProDom: PD003042; HH_signal; 1.
DR SMART: SM00306; Hhntv; 1.
FT NON_TER 1 177
FT SEQUENCE 177 AA; 19739 MW; CA5871626A565F65 CRC64;

Query Match 28.4%; Score 702; DB 11; Length 177;
Best Local Similarity 74.0%; Pred. No. 4, 2e-45;
Matches 131; Conservative 16; Mismatches 30; Indels 0; Gaps 0;

QY 97 LMTORCKDKLNALAIISVNMNPGVKLVNTEGWDGHHSESLHYEGRAVDITTSDBRS 156
DB 1 LMTORCKDKLNALAIISVNMNPGVKLVNTEGWDGHHSESLHYEGRAVDITTSDBRN 60
QY 157 KYGMARLAVEAGFDWYVESKAHICSVKAENSVAASGCGPGSATVHLEGGGKLVK 216
DB 61 KYGMARLAVEAGFDWYVESKAHICSVKAENSVAASGCGPGSATVHLEGGGKLVK 120
QY 217 DLSFGDVLAAADQGRILYSDFTFLDRDDGAKKVFYVETREPRRLTLTAHLLE 273
DB 121 AVKPSGRVLAMGEDGNPTPSDVLIFLDRPNRLRAFOVITDPPRRLLTPAHLLE 177

RESULT 14
Q96699 PRELIMINARY: PRT: 185 AA.
ID 096699;
AC 096699;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Hedgehog protein (Fragment).
OS Junonia coenia (Peacock butterfly) (Precis coenia).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
```

```
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysa;
OC Papilionoidea; Nymphalidae; Nymphalinae; Junonia.
OX NCBITaxID=39708;
RN [1]
RP SEQUENCE FROM N.A.
RA Keys D.N., Lewis D.L., Selegue J.E., Pearson B.J., Goodrich L.V.,
RA Johnson R.L., Gates J., Scott M.P., Carroll S.B.;
RT "Recruitment of a hedgehog regulatory circuit in butterfly eyespot
RT evolution.";
RL Science 0:0-0(1999).
DR EMBL: AF117742; AAD08931.1; -.
DR HSSP: 062226; 1VH.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR001657; SonicHH.
DR Pfam: PF01085; HH_signal; 1.
DR PRINTS: PR00632; SONICHHOG.
DR ProDom: PD003042; HH_signal; 1.
FT NON_TER 1 185
FT SEQUENCE 185 AA; 20745 MW; E9B9F4CE2FAAE662 CRC64;

Query Match 25.1%; Score 619; DB 5; Length 185;
Best Local Similarity 68.2%; Pred. No. 7, 7e-39;
Matches 116; Conservative 21; Mismatches 31; Indels 2; Gaps 2;

QY 58 ASGRYEGKISRNSERKELTPNYPNIIIFKDEENTGADRLMORCKDKLNALAIISVNMN 117
DB 3 ASGRYEGKISRNSERKELTPNYPNIIIFKDEENTGADRLMORCKDKLNALAIISVNMN 62
QY 118 PGVKLVNTEGWDGHHSESLHYEGRAVDITTSDBRSKYGMARLAVEAGFDWYVES 177
DB 63 PGVKLVNTEGWDGHHSESLHYEGRAVDITTSDBRSKYGMARLAVEAGFDWYVES 122
QY 178 KAHICSVKAENSVAASGCGPGSATVHLEGGGKLVNLEGGGKLVNLEGGGKLVN 227
DB 123 KAHICSVKAENSVAASGCGPGSATVHLEGGGKLVNLEGGGKLVNLEGGGKLVN 170

RESULT 15
Q96GN4 PRELIMINARY: PRT: 99 AA.
ID 096GN4;
AC 096GN4;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Sonic hedgehog (Fragment).
OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphi; Galeoidea; Carchariniiformes;
OC Scyllorhinidae; Scyllorhinus.
OX NCBITaxID=7830;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21930458; PubMed=11932743;
RA Tanaka M., Munsterberg A., Anderson W.G., Prescott A.R., Hazon N.,
RA Tickle C.;
RT "Fin development in a cartilaginous fish and the origin of vertebrate
RT limbs.";
RL Nature 416:527-531(2002).
DR EMBL: AF393835; AAM08228.1; -.
FT NON_TER 1 99
FT SEQUENCE 99 AA; 11468 MW; 831864BD834A1A20 CRC64;

Query Match 21.5%; Score 531; DB 13; Length 99;
Best Local Similarity 98.0%; Pred. No. 1, 3e-32;
Matches 97; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 78 PNYPNIIIFKDEENTGADRLMORCKDKLNALAIISVNMNPGVKLVNTEGWDGHHSE 137
DB 1 PNYPNIIIFKDEENTGADRLMORCKDKLNALAIISVNMNPGVKLVNTEGWDGHHSE 60
QY 138 SLHYEGRAVDITTSDBRSKYGMARLAVEAGFDWYVE 176
```

Db 61 SLHREGRAVDITSDRDRKTKGMLARLAVEAGFDWVYFE 99

Search completed: February 20, 2003, 10:11:53
Job time : 32.5535 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:10:28 ; Search time 8.52655 Seconds
(without alignments)
1423.285 Million cell updates/sec

Title: US-09-827-110A-15
Perfect score: 2469
Sequence: 1 MLLARCLLVSLVSLVCS.....GMAVSKSRGAGGAREGA 475

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*
1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB pep.*
7: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB pep.*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB pep.*
10: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB pep.*
11: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB pep.*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2467	99.9	475	8	US-08-900-220C-15
2	2467	99.9	475	9	US-09-883-848A-15
3	2467	99.9	475	10	US-09-021-660A-39
4	2467	99.9	475	10	US-09-151-999-15
5	2407	97.5	462	9	US-09-733-634-14
6	2068	83.8	437	9	US-09-969-520A-1
7	2068	83.8	437	9	US-10-013-310-1
8	2068	83.8	437	9	US-09-733-634-16
9	2068	83.8	437	10	US-09-021-660A-37
10	2065	83.6	437	8	US-08-900-220C-13
11	2065	83.6	437	9	US-09-883-848A-13
12	2065	83.6	437	10	US-09-151-999-13
13	2062	83.5	437	9	US-09-990-046-14
14	2062	83.5	437	9	US-09-969-520A-3
15	2059	83.4	437	9	US-09-969-520A-10
16	2059	83.4	437	9	US-09-969-520A-11
17	2058	83.4	437	9	US-09-969-520A-2
18	2054	83.2	437	9	US-09-969-520A-9
19	2052	83.1	437	9	US-09-969-520A-4

20	2048	82.9	437	9	US-09-969-520A-8	Sequence 8, Appl1
21	2045	82.8	437	9	US-09-969-520A-6	Sequence 6, Appl1
22	2036	82.5	437	9	US-09-969-520A-7	Sequence 7, Appl1
23	2019	81.8	437	9	US-09-969-520A-5	Sequence 5, Appl1
24	1807	73.2	425	8	US-08-900-220C-10	Sequence 10, Appl1
25	1807	73.2	425	9	US-09-883-848A-10	Sequence 10, Appl1
26	1807	73.2	425	10	US-09-021-660A-34	Sequence 34, Appl1
27	1807	73.2	425	10	US-09-151-999-10	Sequence 10, Appl1
28	1506	61.0	418	8	US-09-021-660A-38	Sequence 38, Appl1
29	1467	59.4	418	8	US-08-900-220C-14	Sequence 14, Appl1
30	1467	59.4	418	9	US-09-883-848A-14	Sequence 14, Appl1
31	1467	59.4	418	10	US-09-151-999-14	Sequence 14, Appl1
32	1432	58.0	416	8	US-08-900-220C-18	Sequence 18, Appl1
33	1432	58.0	416	9	US-09-883-848A-18	Sequence 18, Appl1
34	1432	58.0	416	10	US-09-151-999-18	Sequence 18, Appl1
35	1271	51.5	411	8	US-08-900-220C-16	Sequence 16, Appl1
36	1271	51.5	411	9	US-09-883-848A-16	Sequence 16, Appl1
37	1271	51.5	411	10	US-09-151-999-16	Sequence 16, Appl1
38	1266	51.3	411	9	US-09-733-634-28	Sequence 28, Appl1
39	1262	51.1	449	9	US-09-990-046-29	Sequence 29, Appl1
40	1262	51.1	449	9	US-09-733-634-30	Sequence 30, Appl1
41	1234	50.0	411	8	US-08-900-220C-12	Sequence 12, Appl1
42	1234	50.0	411	9	US-09-883-848A-12	Sequence 12, Appl1
43	1234	50.0	411	10	US-09-151-999-12	Sequence 12, Appl1
44	1183	47.9	396	8	US-08-900-220C-11	Sequence 11, Appl1
45	1183	47.9	396	9	US-09-990-046-13	Sequence 13, Appl1

ALIGNMENTS

RESULT 1
US-08-900-220C-15
Sequence 15, Application US/08900220C
Patent No. US20020045206A1
GENERAL INFORMATION:
APPLICANT: Miao, Ningning
Wang, Monica
Mahanthappa, Nagesh K.
Jin, Ping
Jin, Ping
TITLE OF INVENTION: Method of Treating Dopaminergic and GABA-nergic Disorders
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: POLLEY, HONG & ELLIOT LLP
STREET: ONE POST OFFICE SQUARE
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Ascii (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/900,220C
FILING DATE: 24-Jul-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: ONV-044, 01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 475 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-08-900-220C-15

Query Match 99.9%; Score 2467; DB 8; Length 475;
Best Local Similarity 100.0%; Pred. No. 1,4e-192;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MLLARCLLVVSSLVCSGLACGPGRGFGRRHPKLTPLAYKQFIPNVAEKTIGASG 60
   |||||||
DB 1 MLLARCLLVVSSLVCSGLACGPGRGFGRRHPKLTPLAYKQFIPNVAEKTIGASG 60
QY 61 RREGKISRNSERFEKLTLPYNDPIIFKDEENTGADRLMTQROCKDKLNLAISVMNQPGV 120
   |||||||
DB 61 RREGKISRNSERFEKLTLPYNDPIIFKDEENTGADRLMTQROCKDKLNLAISVMNQPGV 120
QY 121 KLRVTEGMDGDHSEESLHYEGRAVDITTSRDRSKTYGMLARLAVEAGFDVYYESKAH 180
   |||||||
DB 121 KLRVTEGMDGDHSEESLHYEGRAVDITTSRDRSKTYGMLARLAVEAGFDVYYESKAH 180
QY 181 HCSYKAENSVAAKSGCGCPGSATVHLEGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
   |||||||
DB 181 HCSYKAENSVAAKSGCGCPGSATVHLEGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
QY 241 FLDRDGAKKVYVETREPRERLLTAHLLFVAPHNDSATGEPEASSGSGPPSGALG 300
   |||||||
DB 241 FLDRDGAKKVYVETREPRERLLTAHLLFVAPHNDSATGEPEASSGSGPPSGALG 300
QY 301 PALFAASRVPRGQRYVVAERDGRRLPAAVHSVTLSEEAAGAVAPLTAOGTILINRYL 360
   |||||||
DB 301 PALFAASRVPRGQRYVVAERDGRRLPAAVHSVTLSEEAAGAVAPLTAOGTILINRYL 360
QY 361 ASCYAVIEEHSNAHRAFAFRLAHALLAALAPARTDRGDSGGGDRGGGGRVALTPAGA 420
   |||||||
DB 361 ASCYAVIEEHSNAHRAFAFRLAHALLAALAPARTDRGDSGGGDRGGGGRVALTPAGA 420
QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLDESEALHPLGMAYKSSXSRGAGGAREGA 475
   |||||||
DB 421 ADAPGAGATAGIHWYSQLLYQIGTWLDESEALHPLGMAYKSSXSRGAGGAREGA 475

```

RESULT 2

US-09-883-848A-15
Sequence 15, Application US/09883848A
Publication No. US20030022819A1
GENERAL INFORMATION:
APPLICANT: Ling, L.
TITLE OF INVENTION: ANGIOGENESIS-MODULATING COMPOSITIONS AND USES
FILE REFERENCE: CIBT-P01-119
CURRENT APPLICATION NUMBER: US/09/883, 848A
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/211, 919
PRIOR FILING DATE: 2000-06-16
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 475
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (463)
OTHER INFORMATION: xaa-unknown amino acid residue
US-09-883-848A-15

Query Match 99.9%; Score 2467; DB 9; Length 475;
Best Local Similarity 100.0%; Pred. No. 1,4e-192;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MLLARCLLVVSSLVCSGLACGPGRGFGRRHPKLTPLAYKQFIPNVAEKTIGASG 60
   |||||||
DB 1 MLLARCLLVVSSLVCSGLACGPGRGFGRRHPKLTPLAYKQFIPNVAEKTIGASG 60

```

```

QY 61 RREGKISRNSERFEKLTLPYNDPIIFKDEENTGADRLMTQROCKDKLNLAISVMNQPGV 120
   |||||||
DB 61 RREGKISRNSERFEKLTLPYNDPIIFKDEENTGADRLMTQROCKDKLNLAISVMNQPGV 120
QY 121 KLRVTEGMDGDHSEESLHYEGRAVDITTSRDRSKTYGMLARLAVEAGFDVYYESKAH 180
   |||||||
DB 121 KLRVTEGMDGDHSEESLHYEGRAVDITTSRDRSKTYGMLARLAVEAGFDVYYESKAH 180
QY 181 HCSYKAENSVAAKSGCGCPGSATVHLEGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
   |||||||
DB 181 HCSYKAENSVAAKSGCGCPGSATVHLEGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
QY 241 FLDRDGAKKVYVETREPRERLLTAHLLFVAPHNDSATGEPEASSGSGPPSGALG 300
   |||||||
DB 241 FLDRDGAKKVYVETREPRERLLTAHLLFVAPHNDSATGEPEASSGSGPPSGALG 300
QY 301 PALFAASRVPRGQRYVVAERDGRRLPAAVHSVTLSEEAAGAVAPLTAOGTILINRYL 360
   |||||||
DB 301 PALFAASRVPRGQRYVVAERDGRRLPAAVHSVTLSEEAAGAVAPLTAOGTILINRYL 360
QY 361 ASCYAVIEEHSNAHRAFAFRLAHALLAALAPARTDRGDSGGGDRGGGGRVALTPAGA 420
   |||||||
DB 361 ASCYAVIEEHSNAHRAFAFRLAHALLAALAPARTDRGDSGGGDRGGGGRVALTPAGA 420
QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLDESEALHPLGMAYKSSXSRGAGGAREGA 475
   |||||||
DB 421 ADAPGAGATAGIHWYSQLLYQIGTWLDESEALHPLGMAYKSSXSRGAGGAREGA 475

```

RESULT 3

US-09-021-660A-39
Sequence 39, Application US/09021660A
Patent No. US20010041668A1
GENERAL INFORMATION:
APPLICANT: Baron, M.
APPLICANT: Belausoff, M.
TITLE OF INVENTION: METHODS FOR MODULATING HEMATOPOIESIS AND VASCULAR
FILE REFERENCE: HUIP-P01-060
CURRENT APPLICATION NUMBER: US/09/021, 660A
PRIOR FILING DATE: 2001-08-27
PRIOR APPLICATION NUMBER: 60/037, 513
PRIOR FILING DATE: 1997-02-10
PRIOR APPLICATION NUMBER: 60/049, 763
PRIOR FILING DATE: 1997-06-16
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 39
LENGTH: 475
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (463)
OTHER INFORMATION: xaa-unknown amino acid
US-09-021-660A-39

Query Match 99.9%; Score 2467; DB 10; Length 475;
Best Local Similarity 100.0%; Pred. No. 1,4e-192;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MLLARCLLVVSSLVCSGLACGPGRGFGRRHPKLTPLAYKQFIPNVAEKTIGASG 60
   |||||||
DB 1 MLLARCLLVVSSLVCSGLACGPGRGFGRRHPKLTPLAYKQFIPNVAEKTIGASG 60
QY 61 RREGKISRNSERFEKLTLPYNDPIIFKDEENTGADRLMTQROCKDKLNLAISVMNQPGV 120
   |||||||
DB 61 RREGKISRNSERFEKLTLPYNDPIIFKDEENTGADRLMTQROCKDKLNLAISVMNQPGV 120
QY 121 KLRVTEGMDGDHSEESLHYEGRAVDITTSRDRSKTYGMLARLAVEAGFDVYYESKAH 180
   |||||||
DB 121 KLRVTEGMDGDHSEESLHYEGRAVDITTSRDRSKTYGMLARLAVEAGFDVYYESKAH 180

```

QY	181	HC5VKAKENSAAKSGGCFPGSATTYHLEBOGGTFLYDLS	PGSRVYLAADDOGLLYSDLT	240
Db	181	HC5VKAKENSAAKSGGCFPGSATTYHLEBOGGTFLYDLS	PGSRVYLAADDOGLLYSDLT	240
QY	241	FLDRDDCAKVFYVYIETREPERRLTLTAAHLTFVAP	PHNDSATGEPEASGSGPPSGALG	300
Db	241	FLDRDDCAKVFYVYIETREPERRLTLTAAHLTFVAP	PHNDSATGEPEASGSGPPSGALG	300
QY	301	PRALFASRVPRGSRVYVVAERDGRRLTLAAVHSYTLSEAA	GAVALPRLTAOCTILINVL	360
Db	301	PRALFASRVPRGSRVYVVAERDGRRLTLAAVHSYTLSEAA	GAVALPRLTAOCTILINVL	360
QY	361	ASCAVYVIEEHSMAHRAFAPERLHAALLAALAPART	DGGDSGGDGGGGGCVALTARGA	420
Db	361	ASCAVYVIEEHSMAHRAFAPERLHAALLAALAPART	DGGDSGGDGGGGGCVALTARGA	420
QY	421	ADAGACATGTHMYSOLLTQIGTWTLLDSEPLAHPLG	MAVKSXSXSGACAGAREGA	475
Db	421	ADAGACATGTHMYSOLLTQIGTWTLLDSEPLAHPLG	MAVKSXSXSGACAGAREGA	475

RESULT 4
 US-09-151-999-15
 Sequence 15, Application US/09151999
 Patent No. US20020151460a1
 GENERAL INFORMATION:
 APPLICANT: Wang, Elizabeth
 TITLE OF INVENTION: REGULATION OF EPITHELIAL TISSUE BY HEDGEHOG-LIKE
 FILE REFERENCE: POLYPEPTIDES, AND FORMULATIONS AND USES RELATED THERETO
 FILE REFERENCE: ONV-031.02
 CURRENT APPLICATION NUMBER: US/09/151,999
 CURRENT FILING DATE: 1998-08-11
 EARLIER APPLICATION NUMBER: 08/955,552
 EARLIER FILING DATE: 1997-10-20
 NUMBER OF SEQ ID NOS: 28
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 15
 LENGTH: 475
 TYPE: PRN
 ORGANISM: Homo sapien Shh
 FEATURE:
 OTHER INFORMATION: Xaa at position 463 is any or unknown amino acid
 US-09-151-999-15

Query Match	99.9%	Score 2467	DB 10	Length 475
Best Local Similarity	100.0%	Pred. No. 1.4e-192		
Matches 475	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	MLLARCILLVYVSSLLVCSGLACGPGRGGRKRRHKLLPTLAYKQFIPTNVAEKTLAGSG	60	
Db	1	MLLARCILLVYVSSLLVCSGLACGPGRGGRKRRHKLLPTLAYKQFIPTNVAEKTLAGSG	60	
Qy	61	RYEKKISRNSEREKELTPNTNPDIITRKDEBNTGADRLMTORCKDXLMLAISVNNQMGV	120	
Db	61	RYEKKISRNSERFKEKELTPNTNPDIITRKDEBNTGADRLMTORCKDKLMLAISVNNQMGV	120	
Qy	121	KLRTYEGMDEGHHSEESLHYEGRAVDITTSDDRSKTYMLRLAVEAGFDWVYYESRAH	180	
Db	121	KLRTYEGMDEGHHSEESLHYEGRAVDITTSDDRSKTYMLRLAVEAGFDWVYYESRAH	180	
Qy	181	IHCYVKAENSVAAKSGGCFPGSATVHLEOGGTFLYVDLSPGGRVLAADDQGLLXSDFLT	240	
Db	181	IHCYVKAENSVAAKSGGCFPGSATVHLEOGGTFLYVDLSPGGRVLAADDQGLLXSDFLT	240	
Qy	241	FLDRDDCAKKVFIYIETREPRERLLITAAHLFLVAPHNDSATGEPDASSGSPGSGALG	300	
Db	241	FLDRDDCAKKVFIYIETREPRERLLITAAHLFLVAPHNDSATGEPDASSGSPGSGALG	300	
Qy	301	PRALFASRVPRGQVYVVAERDDRRLLPAAVHSYVLSLEAGAAAPLTAQGTLLINVL	360	
Db	301	PRALFASRVPRGQVYVVAERDDRRLLPAAVHSYVLSLEAGAAAPLTAQGTLLINVL	360	

Oy 361 ASCAAVIEHSHMARAFEPFLTAHALAALPATDDBGSGGDRGGGGGRVALTARGA 420
 Db 361 ASCAAVIEHSHMARAFEPFLTAHALAALPATDDBGSGGDRGGGGGRVALTARGA 420
 Oy 421 ADARAGAGTAAIHHYSOLLVIGTWLLDSEALHPLGAAVSSSXSRGACGGCARBEA 475
 Db 421 ADARAGAGTAAIHHYSOLLVIGTWLLDSEALHPLGAAVSSSXSRGACGGCARBEA 475

```

RESULT 5
US-09-733-634-14
: Sequence 14, Application US/09733634
: Publication No. US20030013646A1
: GENERAL INFORMATION:
: APPLICANT: Massachusetts General Hospital
: TITLE OF INVENTION: Method to stimulate insulin production by pancreatic b-cells
: FILE REFERENCE: 17633/1240
: CURRENT APPLICATION NUMBER: US/09/733,634
: CURRENT FILING DATE: 2000-12-08
: PRIOR APPLICATION NUMBER: US 60/170,282
: PRIOR FILING DATE: 1999-12-10
: NUMBER OF SEQ ID NOS: 32
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 14
: LENGTH: 462
: TYPE: PR1
: ORGANISM: Homo sapiens
US-09-733-634-14

```

Query Match	97.5%	Score 2407	DB 9	Length 462
Best Local Similarity	100.0%	Pred. NC. 1e-187		
Matches 462; Conservative	0	Mismatches	0	Gaps 0

OY	1	MLARCLLYVSSLLVSSGLAGPGGPGCKRRHPKRLTPYAKOTIPVAEXTTLAGSG	60
Db	1	MLLARCLLYVSSLLVSSGLAGPGGPGGCKRRHPKRLTPYAKOTIPVAEXTTLAGSG	60
OY	61	RYEGKISNSERFELTPPNYNDITFKDEENTGADRLMTORCKOKLNAIAISVNONPGV	120
Db	61	RYEGKISNSERFELTPPNYNDITFKDEENTGADRLMTORCKOKLNAIAISVNONPGV	120
OY	121	KLRYTEGMDDEDHHSSESLHYEGRAVDITTSDDRSRKYGMLARLAVAGEFDWYYESKAH	180
Db	121	KLRYTEGMDDEDHHSSESLHYEGRAVDITTSDDRSRKYGMLARLAVAGEFDWYYESKAH	180
OY	181	IHCYSKAENSVAAKSGGCGFPGSATVHLBOGGTKLVKDISPDQRYLAADDGGRLLYSDFLT	240
Db	181	IHCYSKAENSVAAKSGGCGFPGSATVHLBOGGTKLVKDISPDQRYLAADDGGRLLYSDFLT	240
OY	241	FLDRDDGAKKFFVYIETFEPEPERLLTNAHLLEFAPPHNDSTGTGPEKSSSGSPSGGALG	300
Db	241	FLDRDDGAKKFFVYIETFEPEPERLLTNAHLLEFAPPHNDSTGTGPEKSSSGSPSGGALG	300
OY	301	PRALFASVRRGORYVVAEEDSGRRLLPAAVHSVTLSSEAAAGYAPLTNOGTLLINRYL	360
Db	301	PRALFASVRRGORYVVAEEDSGRRLLPAAVHSVTLSSEAAAGYAPLTNOGTLLINRYL	360
OY	361	ASCVAVIEHSHMARAFAPFLAHLAALLAALPARTDGCOSGGCDRGGCGGRVALTPAGA	420
Db	361	ASCVAVIEHSHMARAFAPFLAHLAALLAALPARTDGCOSGGCDRGGCGGRVALTPAGA	420
OY	421	ADAPGAGTAGIHWYSOLLYOIGTWLLDSEALHPLGNAVSS 462	
Db	421	ADAPGAGTAGIHWYSOLLYOIGTWLLDSEALHPLGNAVSS 462	
RESULT 6			
US-09-969-520A-1			
; Sequence 1, Application US/09969520A			
; Patent No. US20020177163A1			
; GENERAL INFORMATION:			
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE			
; APPLICANT: BEACHY, Phillip A.			

```

; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED
; FILE REFERENCE: JH01670-1
; CURRENT APPLICATION NUMBER: US/09/969,520A
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/235,153
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-969-520A-1

Query Match      83.8%; Score 2068; DB 9; Length 437;
Best Local Similarity 87.4%; Pred. No. 3,1e-160;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 2 LLLARCFVLVIASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGASG 61
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
QY 61 RYEGKISRNSERFKELTPYNPDIIFKDEENTGADRLMTQRCCKDLNALAISVMNQPGV 120
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 62 RYEGKISRNSERFKELTPYNPDIIFKDEENTGADRLMTQRCCKDLNALAISVMNQPGV 121
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
QY 121 KLRVTEGMDEGDHSEESLHYEGRAVDITTSRDRSKYGMLARLAVEAGFDVYYESKAH 180
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 122 KLRVTEGMDEGDHSEESLHYEGRAVDITTSRDRSKYGMLARLAVEAGFDVYYESKAH 181
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
QY 181 IHCYKAENSVAAKSGCGCPGSAIYHLEGGTKLVKDLSPGRVLAADQGRILYSDFLT 240
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 182 IHCYKAENSVAAKSGCGCPGSAIYHLEGGTKLVKDLSPGRVLAADQGRILYSDFLT 241
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
QY 241 FLDRDGAKKVYVYETREPRERLLTAHLLFVAPHNDSATGEPPASSGSGPPSGALG 300
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 242 FLDRDGAKKVYVYETREPRERLLTAHLLFVAPHNDSATGEPPASSGSGPPSGALG 300
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
QY 301 PALFASRVRRPGQRYVVAERDGRRLPAAVHSVTLSEBAAGAVAPLTAOSTIILINRVL 360
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 287 PALFASRVRRPGQRYVVAERDGRRLPAAVHSVTLSEBAAGAVAPLTAOSTIILINRVL 360
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
QY 361 ASCYAVIEHSHVAHAFAPFRLAHLAALAPARTDRGDSGCGRGGGVALTAPGA 420
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 347 ASCYAVIEHSHVAHAFAPFRLAHLAALAPARTD-----GGCGGSIP-AAQSA 395
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
QY 421 AAPAGAGATAGIHWYSQLLYOIGTWLDSSEALHPGLMAVKSS 462
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 396 TEARGAEPTAGIHWYSQLLYHIGTWLDSSEALHPGLMAVKSS 437
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||

RESULT 7
US-10-013-310-1
; Sequence 1, Application US/10013310
; Publication No. US20020192216A1
; GENERAL INFORMATION:
; APPLICANT: Lamb, Jonathan Robert
; APPLICANT: Hoyne, Gerard Francis
; APPLICANT: Dallman, Margaret Jane
; TITLE OF INVENTION: Therapeutic Use
; FILE REFERENCE: 674525-2003
; CURRENT APPLICATION NUMBER: US/10/013,310
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: PCT/G800/02191
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: UK 9913350.6
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: UK 9921953.7
; PRIOR FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 437
```

```

; TYPE: PRT
; ORGANISM: House Mouse
US-10-013-310-1

Query Match      83.8%; Score 2068; DB 9; Length 437;
Best Local Similarity 87.4%; Pred. No. 3,1e-160;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 2 LLLARCFVLVIASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGASG 61
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
QY 61 RYEGKISRNSERFKELTPYNPDIIFKDEENTGADRLMTQRCCKDLNALAISVMNQPGV 120
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 62 RYEGKISRNSERFKELTPYNPDIIFKDEENTGADRLMTQRCCKDLNALAISVMNQPGV 121
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
QY 121 KLRVTEGMDEGDHSEESLHYEGRAVDITTSRDRSKYGMLARLAVEAGFDVYYESKAH 180
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 122 KLRVTEGMDEGDHSEESLHYEGRAVDITTSRDRSKYGMLARLAVEAGFDVYYESKAH 181
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
QY 181 IHCYKAENSVAAKSGCGCPGSAIYHLEGGTKLVKDLSPGRVLAADQGRILYSDFLT 240
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 182 IHCYKAENSVAAKSGCGCPGSAIYHLEGGTKLVKDLSPGRVLAADQGRILYSDFLT 241
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
QY 241 FLDRDGAKKVYVYETREPRERLLTAHLLFVAPHNDSATGEPPASSGSGPPSGALG 300
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 242 FLDRDGAKKVYVYETREPRERLLTAHLLFVAPHNDSATGEPPASSGSGPPSGALG 300
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
QY 301 PALFASRVRRPGQRYVVAERDGRRLPAAVHSVTLSEBAAGAVAPLTAOSTIILINRVL 360
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 287 PALFASRVRRPGQRYVVAERDGRRLPAAVHSVTLSEBAAGAVAPLTAOSTIILINRVL 360
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
QY 361 ASCYAVIEHSHVAHAFAPFRLAHLAALAPARTDRGDSGCGRGGGVALTAPGA 420
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 347 ASCYAVIEHSHVAHAFAPFRLAHLAALAPARTD-----GGCGGSIP-AAQSA 395
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
QY 421 AAPAGAGATAGIHWYSQLLYOIGTWLDSSEALHPGLMAVKSS 462
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 396 TEARGAEPTAGIHWYSQLLYHIGTWLDSSEALHPGLMAVKSS 437
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||

RESULT 8
US-09-733-634-16
; Sequence 16, Application US/09733634
; Publication No. US20030013646A1
; GENERAL INFORMATION:
; APPLICANT: Massachusetts General Hospital
; TITLE OF INVENTION: Method to stimulate insulin production by pancreatic b-cells
; FILE REFERENCE: 17633/1240
; CURRENT APPLICATION NUMBER: US/09/733,634
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/170,282
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-733-634-16

Query Match      83.8%; Score 2068; DB 9; Length 437;
Best Local Similarity 87.4%; Pred. No. 3,1e-160;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 2 LLLARCFVLVIASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGASG 61
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
QY 61 RYEGKISRNSERFKELTPYNPDIIFKDEENTGADRLMTQRCCKDLNALAISVMNQPGV 120
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
DB 62 RYEGKISRNSERFKELTPYNPDIIFKDEENTGADRLMTQRCCKDLNALAISVMNQPGV 121
   :||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:||||| :|:|||||
```


Db 122 RLRTGEMDECHSESLHYEGRAVDITTSRDRSKYGMRLRLAVEAGFDWVYESKAH 181
QY 181 IHCSYKAENSVAAKSGGCPGSATVHLEOGGCTKLVKDLSPGDRVLAADOGRLTSDFLT 240
Db 182 IHCSYKAENSVAAKSGGCPGSATVHLEOGGCTKLVKDLSPGDRVLAADOGRLTSDFLT 241
QY 241 FLDRDGAKKVYVYETREPRERLLTAANHLFVAPHNDSATGEPBASSGSGPPSGALG 300
Db 242 FLDRDGAKKVYVYETREPRERLLTAANHLFVAPHND-----SGPPPG 286
QY 301 PALFASRYRPGQRYVVAERDGRRLRLPAVHSTLSEAGAAYAPLTAOGTILINRL 360
Db 287 PALFASRYRPGQRYVVAERDGRRLRLPAVHSTLSEAGAAYAPLTAOGTILINRL 346
QY 361 ASCYAVIEHSHAHRAFAFRLAHLAALAPARTDRGDSGGGGRGVALTAPGA 420
Db 347 ASCYAVIEHSHAHRAFAFRLAHLAALAPARTD-----GGGGGSIIP-AAQSA 395
QY 421 ADAPGAGATAGIHWYSQLYQIGTWLDSSEALHPLGMAYKSS 462
Db 396 TEARGAEPTAGIHWYSQLYHIGTWLDSETHMPLGMAYKSS 437

RESULT 11

US-09-883-848A-13
; Sequence 13, Application US/09883848A
; Publication No. US20030022819A1
; GENERAL INFORMATION:
; APPLICANT: Ling, L.
; APPLICANT: Sanicola-Nadel, M.
; TITLE OF INVENTION: ANGIOGENESIS-MODULATING COMPOSITIONS AND USES
; FILE REFERENCE: CIBT-P01-119
; CURRENT APPLICATION NUMBER: US/09/883, 848A
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/211, 919
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-883-848A-13

Query Match 83.6%; Score 2065; DB 9; Length 437;

Best Local Similarity 87.2%; Pred. No. 5,4e-160;
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

QY 1 MILLARCLLVYSSLLVCSGLACGPGFGRRRHPKLTPLAYKOTIPNVAEKTILGASG 60
Db 2 LLLLRKCFVLIASSLLVCPGLACGPGFGRRRHPKLTPLAYKOTIPNVAEKTILGASG 61
QY 61 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVNMOMPGV 120
Db 62 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVNMOMPGV 121
QY 121 KLRTVEGMEDECHSESLHYEGRAVDITTSRDRSKYGMRLRLAVEAGFDWVYESKAH 180
Db 122 RLRTGEMDECHSESLHYEGRAVDITTSRDRSKYGMRLRLAVEAGFDWVYESKAH 181
QY 181 IHCSYKAENSVAAKSGGCPGSATVHLEOGGCTKLVKDLSPGDRVLAADOGRLTSDFLT 240
Db 182 IHCSYKAENSVAAKSGGCPGSATVHLEOGGCTKLVKDLSPGDRVLAADOGRLTSDFLT 241
QY 241 FLDRDGAKKVYVYETREPRERLLTAANHLFVAPHNDSATGEPBASSGSGPPSGALG 300
Db 242 FLDRDGAKKVYVYETREPRERLLTAANHLFVAPHND-----SGPPPG 286
QY 301 PALFASRYRPGQRYVVAERDGRRLRLPAVHSTLSEAGAAYAPLTAOGTILINRL 360
Db 287 PALFASRYRPGQRYVVAERDGRRLRLPAVHSTLSEAGAAYAPLTAOGTILINRL 346
QY 361 ASCYAVIEHSHAHRAFAFRLAHLAALAPARTDRGDSGGGGRGVALTAPGA 420

Db 347 ASCYAVIEHSHAHRAFAFRLAHLAALAPARTD-----GGGGGSIIP-AAQSA 395
QY 421 ADAPGAGATAGIHWYSQLYQIGTWLDSSEALHPLGMAYKSS 462
Db 396 TEARGAEPTAGIHWYSQLYHIGTWLDSETHMPLGMAYKSS 437

RESULT 12

US-09-151-999-13
; Sequence 13, Application US/09151999
; Patent No. US20020151460A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Elizabeth
; TITLE OF INVENTION: REGULATION OF EPITHELIAL TISSUE BY HEDGEHOG-LIKE
; FILE REFERENCE: ONV-031.02
; CURRENT APPLICATION NUMBER: US/09/151, 999
; PRIOR FILING DATE: 1998-08-11
; EARLIER APPLICATION NUMBER: 08/955, 552
; PRIOR FILING DATE: 1997-10-20
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 437
; TYPE: PRT
; ORGANISM: murine Shh
US-09-151-999-13

Query Match 83.6%; Score 2065; DB 10; Length 437;
Best Local Similarity 87.2%; Pred. No. 5,4e-160;
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

QY 1 MILLARCLLVYSSLLVCSGLACGPGFGRRRHPKLTPLAYKOTIPNVAEKTILGASG 60
Db 2 LLLLRKCFVLIASSLLVCPGLACGPGFGRRRHPKLTPLAYKOTIPNVAEKTILGASG 61
QY 61 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVNMOMPGV 120
Db 62 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVNMOMPGV 121
QY 121 KLRTVEGMEDECHSESLHYEGRAVDITTSRDRSKYGMRLRLAVEAGFDWVYESKAH 180
Db 122 RLRTGEMDECHSESLHYEGRAVDITTSRDRSKYGMRLRLAVEAGFDWVYESKAH 181
QY 181 IHCSYKAENSVAAKSGGCPGSATVHLEOGGCTKLVKDLSPGDRVLAADOGRLTSDFLT 240
Db 182 IHCSYKAENSVAAKSGGCPGSATVHLEOGGCTKLVKDLSPGDRVLAADOGRLTSDFLT 241
QY 241 FLDRDGAKKVYVYETREPRERLLTAANHLFVAPHNDSATGEPBASSGSGPPSGALG 300
Db 242 FLDRDGAKKVYVYETREPRERLLTAANHLFVAPHND-----SGPPPG 286
QY 301 PALFASRYRPGQRYVVAERDGRRLRLPAVHSTLSEAGAAYAPLTAOGTILINRL 360
Db 287 PALFASRYRPGQRYVVAERDGRRLRLPAVHSTLSEAGAAYAPLTAOGTILINRL 346
QY 361 ASCYAVIEHSHAHRAFAFRLAHLAALAPARTDRGDSGGGGRGVALTAPGA 420
Db 347 ASCYAVIEHSHAHRAFAFRLAHLAALAPARTD-----GGGGGSIIP-AAQSA 395
QY 421 ADAPGAGATAGIHWYSQLYQIGTWLDSSEALHPLGMAYKSS 462
Db 396 TEARGAEPTAGIHWYSQLYHIGTWLDSETHMPLGMAYKSS 437

RESULT 13

US-09-990-046-14
; Sequence 14, Application US/09990046
; Patent No. US20020156245A1
; GENERAL INFORMATION:
; APPLICANT: de Sauvage, Frederic
; APPLICANT: Carpenter, David A.

;; TITLE OF INVENTION: Patched-2
;; FILE REFERENCE: P1405RI
;; CURRENT APPLICATION NUMBER: US/09/990,046
;; CURRENT FILING DATE: 2001-11-20
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/293,505
;; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-15
;; NUMBER OF SEQ ID NOS: 32
;; SEQ ID NO 14
;; LENGTH: 437
;; TYPE: PRT
;; ORGANISM: Mus musculus
US-09-990-046-14

Query Match 83.5%; Score 2062; DB 9; Length 437;
Best Local Similarity 87.0%; Pred. No. 9, 4e-160;
Matches 402; Conservative 10; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLVVSSLLVCGSLACGPGRGKRRHKKLPLAYKOFIPVNAEKTIGASC 60
DB 2 LLLARCLFVLVSSLLVCGSLACGPGRGKRRHKKLPLAYKOFIPVNAEKTIGASC 61
QY 61 RYEGKISRNSERKELTPNNPDIIKFDEENTGADRLMTORCKDKLALISVNMOPGV 120
DB 62 RYEGKISRNSERKELTPNNPDIIKFDEENTGADRLMTORCKDKLALISVNMOPGV 121
QY 121 KLRVTEGMDGDHSESLHYEGRAVDITTSDBRSKYGLARLAVAGEFDWYYESKAH 180
DB 122 KLRVTEGMDGDHSESLHYEGRAVDITTSDBRSKYGLARLAVAGEFDWYYESKAH 181
QY 181 IHCSYKAENSVAAKSGGCFPGSATVHLEOGCTKLVKDLRPGDRVLAADDGRLLYSDFLT 240
DB 182 IHCSYKAENSVAAKSGGCFPGSATVHLEOGCTKLVKDLRPGDRVLAADDGRLLYSDFLT 241
QY 241 FLDRDGAKKVFYVETREPERELLTAHLLEFVAPHNDSATGEPASSGSGPPSGALG 300
DB 242 FLDRDEAKKVFYVETLEPERELLTAHLLEFVAPHNDSATGEPASSGSGPPSGALG 300
QY 301 PRALFASRVPRGQRYVVAERDGRLLPAAVHSVTLSEBAAGAYAPLTAQGTILINRVL 360
DB 287 PSALFASRVPRGQRYVVAERDGRLLPAAVHSVTLSEBAAGAYAPLTAHGTILINRVL 346
QY 361 ASCYAVTEESHMAHRAFPRLAHALLAALAPARTDRCGSGGCGGGRGVALTPGA 420
DB 347 ASCYAVTEESHMAHRAFPRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395
QY 421 ADAPGAGATAGIHWSOLYQIGTWLDSALPLPLGMAVNSS 462
DB 396 TEARGAEPITAGIHWSOLYHIGTWLDSALPLPLGMAVNSS 437

RESULT 14
US-09-969-520A-3
;; Sequence 3, Application US/09969520A
;; Patent No. US20020177163A1
;; GENERAL INFORMATION:
;; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
;; APPLICANT: BEACHY, Philip A.
;; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED
;; FILE REFERENCE: JHU1670-1
;; CURRENT APPLICATION NUMBER: US/09/969,520A
;; CURRENT FILING DATE: 2002-06-04
;; PRIOR APPLICATION NUMBER: US 60/235,153
;; PRIOR FILING DATE: 2000-09-22
;; NUMBER OF SEQ ID NOS: 17
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 3
;; LENGTH: 437
;; TYPE: PRT
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Altered sonic hedgehog protein
US-09-969-520A-3

Query Match 83.5%; Score 2062; DB 9; Length 437;
Best Local Similarity 87.2%; Pred. No. 9, 4e-159;
Matches 403; Conservative 8; Mismatches 25; Indels 26; Gaps 3;

QY 1 MLLARCLLVVSSLLVCGSLACGPGRGKRRHKKLPLAYKOFIPVNAEKTIGASC 60
DB 2 LLLARCLFVLVSSLLVCGSLACGPGRGKRRHKKLPLAYKOFIPVNAEKTIGASC 61
QY 61 RYEGKISRNSERKELTPNNPDIIKFDEENTGADRLMTORCKDKLALISVNMOPGV 120
DB 62 RYEGKISRNSERKELTPNNPDIIKFDEENTGADRLMTORCKDKLALISVNMOPGV 121
QY 121 KLRVTEGMDGDHSESLHYEGRAVDITTSDBRSKYGLARLAVAGEFDWYYESKAH 180
DB 122 KLRVTEGMDGDHSESLHYEGRAVDITTSDBRSKYGLARLAVAGEFDWYYESKAH 181
QY 181 IHCSYKAENSVAAKSGGCFPGSATVHLEOGCTKLVKDLRPGDRVLAADDGRLLYSDFLT 240
DB 182 IHCSYKAENSVAAKSGGCFPGSATVHLEOGCTKLVKDLRPGDRVLAADDGRLLYSDFLT 241
QY 241 FLDRDGAKKVFYVETREPERELLTAHLLEFVAPHNDSATGEPASSGSGPPSGALG 300
DB 242 FLDRDEAKKVFYVETLEPERELLTAHLLEFVAPHNDSATGEPASSGSGPPSGALG 300
QY 301 PRALFASRVPRGQRYVVAERDGRLLPAAVHSVTLSEBAAGAYAPLTAQGTILINRVL 360
DB 287 PSALFASRVPRGQRYVVAERDGRLLPAAVHSVTLSEBAAGAYAPLTAHGTILINRVL 346
QY 361 ASCYAVTEESHMAHRAFPRLAHALLAALAPARTDRCGSGGCGGGRGVALTPGA 420
DB 347 ASCYAVTEESHMAHRAFPRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395
QY 421 ADAPGAGATAGIHWSOLYQIGTWLDSALPLPLGMAVNSS 462
DB 396 TEARGAEPITAGIHWSOLYHIGTWLDSALPLPLGMAVNSS 437

RESULT 15
US-09-969-520A-10
;; Sequence 10, Application US/09969520A
;; Patent No. US20020177163A1
;; GENERAL INFORMATION:
;; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
;; APPLICANT: BEACHY, Philip A.
;; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHE
;; FILE REFERENCE: JHU1670-1
;; CURRENT APPLICATION NUMBER: US/09/969,520A
;; CURRENT FILING DATE: 2002-06-04
;; PRIOR APPLICATION NUMBER: US 60/235,153
;; PRIOR FILING DATE: 2000-09-22
;; NUMBER OF SEQ ID NOS: 17
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 10
;; LENGTH: 437
;; TYPE: PRT
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Altered sonic hedgehog protein
US-09-969-520A-10

Query Match 83.4%; Score 2059; DB 9; Length 437;
Best Local Similarity 87.0%; Pred. No. 1, 6e-159;
Matches 402; Conservative 9; Mismatches 25; Indels 26; Gaps 3;

QY 1 MLLARCLLVVSSLLVCGSLACGPGRGKRRHKKLPLAYKOFIPVNAEKTIGASC 60
DB 2 LLLARCLFVLVSSLLVCGSLACGPGRGKRRHKKLPLAYKOFIPVNAEKTIGASC 61
QY 61 RYEGKISRNSERKELTPNNPDIIKFDEENTGADRLMTORCKDKLALISVNMOPGV 120
DB 62 RYEGKISRNSERKELTPNNPDIIKFDEENTGADRLMTORCKDKLALISVNMOPGV 121
QY 121 KLRVTEGMDGDHSESLHYEGRAVDITTSDBRSKYGLARLAVAGEFDWYYESKAH 180

```
Db 122 KLRTEGWEDEGHSEESLHYEGRAVDITTSADRAKYGMLARLAVEAGFDMVYYESKAH 181
Oy 181 IHCSYKAENSVAKSGGCEPGSATVHLEQGGTKLVKDLSPGDRVLAADDOGRLYSDFLT 240
Db 182 IHCSYKAENSVAKSGGCEPGSATVHLEQGGTKLVKDLRPGDRVLAADDOGRLYSDFLT 241
Oy 241 FLDRDDGAKKVFYIETREPRERLLTAHLLFVA PHNDSATGEPEASSGSGPPSGALG 300
Db 242 FLDRDGAKKVFYIETLEPRERLLTAHLLFVA PHND-----SGPTPG 286
Oy 301 PRALFASRVRRPGQRYVVAERDGRRLPRAVHSTLSEAGAYAPLTAOCTIILINRYL 360
Db 287 PSALFASRVRRPGQRYVVAERDGRRLPRAVHSTLSEAGAYAPLTAOCTIILINRYL 346
Oy 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRCGSGGDRGGRVALTAPCA 420
Db 347 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTD-----GGGGGSIP-AAOSA 395
Oy 421 ADAPGAGATAGIHWYSQLLYQIGTWLDESEALHPLGMAYKSS 462
Db 396 TEARGAEPTAGIHWYSQLLYHIGTWLDESETHPLGMAYKSS 437
```

Search completed: February 20, 2003, 10:14:43
Job time : 9.52655 secs